SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Ainley, Michael
Armstrong, Katherine
Belmar, Scott
Folkerts, Otto
Hopkins, Nicole
Menke, Michael A.
Pareddy, Dayakar
Petolino, Joseph F.
Smith, Kelley
Woosley, Aaron

(ii) TITLE OF INVENTION: Regulatory Sequences for Transgenic Plants

(iii) NUMBER OF SEQUENCES: 59

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: DowAgroSciences LLC
(B) STREET: 9330 Zionsville Road
(C) CITY: Indianapolis
(D) STATE: Indiana
(E) COUNTRY: USA
(F) ZIP: 46268

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Kraus, Eric J.

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 317 337 5110
(B) TELEFAX: 317 337 4847

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6550 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 4201..4425
(D) OTHER INFORMATION: /product= "Peroxidase"

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 4426..5058

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 5059..5250

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 5251..5382

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 5383..5548

(ix) FEATURE:
(A) NAME/KEY: intron
(B) LOCATION: 5549..5649

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 5650..6065

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: join(4201..4425, 5059..5250, 5383..5547, 5649
..6068)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCATGGCCAG TTGCCGGTGG AGCAGGTAAA AACACCGTAG CGTAGCAGCC AGGCGGAAGC 60

AGACGCACAG CACAGGTTGG TTATGATAGT CAGCCGGGCC ACATGTGTGT AGTTGGTACA 120

CTGATACGCT TACACTGTCT CTCCTTTCTT TTTTATTTGT CACCTTTGGT CGAGCTTACA 180

TAATTGTGTG ACTAAAAAAA GGTCACTTCA TTCAGAAATT TAGGGTTGTG GGAATTTTGG 240

ATTTTATTGT GTCTGTATAG AGTAGCTATA GCTAGCTAGC TAGATGTGAT GTTAATAATT 300

ATGACGATGA GATTGGCCCG CTTGGCCGCT TGCATTGTCT CCCTAGCTCA ATAATGTTTT 360

GAGTTTGTCT TGCCTTTCTT TCAGCTCTAA CAAATTGGAG TAGGGATGAC TGAGATACAT 420

ATATAAAAGC GAAAACCGCT GCTCTCTGTT AATTATTGCA CATCACACAT AGGCCAAGCC 480

TTAAGGACAA TCAACTAAGG ATGGTAATAA CTAAGGCTAG TGAGGTCGAA CTAGGGATGT 540

TAATATACTC TAGATTTTAG ACTATAAAAT TTAAGGATCG AATCAGATTA GTATCGAACT 600

ATATTTATAT TCATTTCTAA ACTAAATTAA TTAAGCACCC TAAATTATTG TGATGAAGAG 660

ACATTTCGAT CGTGATCCAT TATTACTCCT TGGTCAAACT AATCTCGTTT TATGTCACTA 720

TTTCATCATC TTTTTTGCGA ACGGGTTTAT AGCCCGTGTT CCATTATGAG GACATGAACG 780

GTTTAAACAA AGTTACATAT CATCCCAGCT AGCTACCTAG ATTGGAAGCA TGGGTTCGGT 840

ATATATATAT AGTTTATATA TTTGGTATAT ATATATATAT ATATATATAT ATATATATAT 900

CACACGTCAG CTTATATTAC GTAAAGTGGG GTTAGTTTTC AAGAAGCGTG GGACCAGTCA 960

CCTCTGCAGT CTGACCTTGG CTTCAGCTTC GACAGCAAAC AGTCATCTCT TGGAAGCTAA 1020

GGACAGTCTC CAACAGTCAA CAAAGCAGCG GTCTGCTTGT AGTTCTCCCT TGCACGACCA 1080

GCTATATCTA GCATCATAAC AACGGTAAGA TCATCTCTAG CACGACAAAC TTAGTTTAAT 1140

TAATTATGTC TAATCCGTTG TTGTTAGCTT AAACTTTCTA GCCTCCTATG CTAAGAGAGT 1200

TCTCTAGTTC TACTCAGGTG GATTGATATA TAAATTGGGA ATCTTCTAGG CGTCACAAGG 1260

TATGGTACAC ATCAATCAAT GAACGGACAA AGCAACGGTA AGATCCGACC CAGTAAAAGT 1320

AATAGCGTTA GGGCATGTAC AACCTAGACA CTGATGCACA GTACTCCAAG TATAAGACAC 1380

AACTAAAACA CAACATAATA ATACAGTGGT TATATCTAAA ACATGTGTCT TACCATATTC 1440

ATTGTACCAA TTAGAACATT TAATAAATTA AAGTGACCAA TCAGCTAGCC TCCTGTCTCG 1500

AACATAGAGC TAAGACATTG TGTCTTCGTC AAGATACATG TCTTAAGTTT TTTTATATTC 1560

ACTCCCAAAG ACACACTCTA AGACACAACG TAACACACCC ATTGTACATG CTCTTAACCT 1620

AAGTTATCAT GGATGACCAC GCGTGGCAAT TAAAAAAATA ATTTTTGCCT CCTAAAACCT 1680

CTTTCTTAAT TGGTTCTTGC TTGCAAATCA CCAGCGAACC CATATGAAAG GATGCTCAAA 1740

ATCTGGCCAC CGCATCAGGG TTGGTGAATG CAACGTAAAA AATAATGCAT AAATCAGCTC 1800

TCTGATCAGT TATATAATCG TGCCTTTTAA TTATTCATGC CAGCTTTATC TGACTCACGA 1860

AATCATTGAT AAATTATTCC TCAGCTGTAT TAGAAAGAGC AGTGTTGTTT AACTTGGAAA 1920

GTGATGTGGA AGCGTGTGAT TGCGGTTGAG CTTGTATAGG AGTAAAATGA GGAACAGTAG 1980

GAAAATAATT TTTTCGGATT AAAACCGGTT GTTTGGACTG CGGCAGATAC AATTCATAGA 2040

GATAAAAACA CCGTAGAAGT ATTAGAAGCC GATAAAGATT AAACCCAAAT GAACGAACAG 2100

GCTAAACAAA TCCGGCGCCT CAAAAGTCAA GAGCAGGTAC TGGGCTGTCT TGCACACGTC 2160

GCTTTTTGTC TCCCCCTGGC CCCTGGGTGA GAGTAGTAGG GATGCTAAAG TTTGCTTTCT 2220

CTTTTTGAGG CATGTGATAG GCTCTTGTTA GTTGCTAGGG CTATGTTTAT AATATTTGCG 2280

CTTTTACCTA TGTACGTAAG AACCGGATGG AATAATGCTA TGCAGGAACC AATTATGTTT 2340

GGTCGAAATA TATAGTGACC TATCATAATG TTATCCCTGT TCATGTACCT AGGTGGCTAA 2400

TGATATACGG CATATGAATA CAGTAATCAT CCAAGCACGT AAAAACTCGC TAGACGTTTA 2460

TGCCTGCTAG CCTGCTGGGT GTGTAGACTG GAGTACTGGA CAAACATCGC AATACAGAGG 2520

TACAGTATTT GTCTAGACAA TGATATACAT AGATAAAAAC CACTGTTGTA ACTTGTAAGC 2580

CACTAGCTCA CGTTCTCCAT GAGCTCTTCT CTCTGCTGTT TCTTCCTCTG CTAACTGCGT 2640

TATGATATGA CGTCGTATAA ATAATCTCAC AATACTTCCT TATTTTCAGC ATGGCCTCTT 2700

TTATGTTTAT TTAACAGTAG CAACCAACGC CGCTCGATGT TTCCTTCAAG AAACGGCCAC 2760

TCACTATGTG GTGTGCAGAA GAACAAATGT AAGCAGCTCC TACAGGTACC AGTAGTCATG 2820

TCAGTGTGGA AGCTTTCCAA CCAACGCCTC CTTCGAGGAA CCTGGTCGTG CTGACATGAA 2880

TGTAGGCCAT GCAAGCACAA GCACCTAACG CGAATCATCA CGACGCGCCG TGTACTGGGC 2940

GTTGGTACAT CACACCCCGC GTTTGACCTG ATCGGAAGCA TGCGTGTGTG TTGGCTGCAG 3000

GACCGGCTAT AGGTTTCCTG CATTGGACAG CAGAAGCCAG TCATGTTAGG CACTCACGCG 3060

CTCCTGCCGT TTGATGAATC ATCCGGTCTT TCGTATTGAT CACTAGTTCA CTACGCTGAT 3120

ATAGCAAATT TTAAGATGTG AAACCACGAG ACGAGCGATA AATCTTAGAC GTTACCTATC 3180

CATATGAAGC TTGTGCGAAA AAAAGGCGTG CCGCTGTAGC ATCATTCGTA TACACTTTTG 3240

TCCCCAAAGA CAGGGATACG AATCCATGCT CGACAGAACC CTCCCTTCCC TGCAGATAAC 3300

GACACTTAAG TATAACAAAA GTAGTTGGAT TATTTCAGAA GCAAAATCTC ACTTTTCGCT 3360

GGCCTTTTTG TACTTTGGTT ACTTGAGTTC AGACAGTGTA TGCTATATTG TCATGTGCTG 3420

CGTAAGGTTT AAATATGGTT CGACAAATAT ATCAGTATAT CACTACTTTG TTATGGGTGG 3480

GGCCTAGCAC AAACTTGATA CAGCTAGGAT AAAGTTAGAA CGATGACTGA TCTACTGTAA 3540

AGCGACACCT GTCCTGTTAT GGTAGTTTAA GTCCATTCCT GGACGACTCC AGATCCAGGA 3600

TATGATGCTG TTACATAATG CGATTGTTCA CAATAAAATT GCATGATGTT CTTCTACTCT 3660

TTAGGCAGTT TTGTTCAACA GGCAAGTTGC ATAATGCATG TGCATATATG AGCAGCATAA 3720

TCATCAATTA ATCATAGGTT CGTCATTTTA GTTTCACTCC TTCACATTAT TCCAGCCCTT 3780

GAAGAAAAAT GTAGCAGTGC TTGCTGTTTA ATAAGTGGCA GAGCTGTTTT CACTCCACCT 3840

ACGCTTGTCT AGGACCAAAA TTTTAATCTG TCACTTTGAG CTAAAACTGA AGCACCAAAC 3900

CGCTACAAAA GAACGTAGGA GCTGAATTGT AACTTGATGG GATTACTATA GCAGTTGCTA 3960

CAGTTCTAGC TAGCTACCTT ATTCTATACG CATCACCCTA ACAACCCGGC TGACTGCTGC 4020

```
ATCTGACCCC ACCGTCCCCT GCTCCAAACC AACTCTCCTT TCCTTGCATG CACTACACCC    4080

ACTTCCTGCA GCTATATATA CCACCATATG CCCATCTTAT GAAACCATCC ACAAGAGGAG    4140

AAGAAACAAT CAACCAGCAA CACTCTTCTC TTATAACATA GTACAGCGAA GGTAACTCAC    4200

ATG GCA ACT TCC ATG GGT TGT CTC GTC TTG CTC TGC CTT GTT TCT TCT      4248
Met Ala Thr Ser Met Gly Cys Leu Val Leu Leu Cys Leu Val Ser Ser
  1               5                  10                  15

CTC CTT CCC AGT GCC GTC CTT GGC CAC CCA TGG GGT GGC TTG TTC CCA      4296
Leu Leu Pro Ser Ala Val Leu Gly His Pro Trp Gly Gly Leu Phe Pro
            20                  25                  30

CAG TTC TAT GAC CAT TCG TGC CCC AAG GCG AAG GAG ATT GTG CAG TCC      4344
Gln Phe Tyr Asp His Ser Cys Pro Lys Ala Lys Glu Ile Val Gln Ser
        35                  40                  45

ATT GTG GCA CAG GCT GTG GCC AAG GAG ACC AGG ATG GCG GCA TCT TTA      4392
Ile Val Ala Gln Ala Val Ala Lys Glu Thr Arg Met Ala Ala Ser Leu
    50                  55                  60

GTC AGA CTG CAT TTC CAT GAC TGC TTT GTC AAG GTTCAATTCT GCTTCCTCTG    4445
Val Arg Leu His Phe His Asp Cys Phe Val Lys
65                  70                  75

TTATGTTCTT TATATTACAT GCTCTGACAA AGCTATAAAG CTTGATACTG CAGTATAATA    4505

TAACAAGTTA GCTACACAAG TTTTGTACTT CAAGTCTTTT AACTATATGT TGGTGCAATA    4565

AGATTATGAG TAATCCATAT GAAGGTGTTG CAAGAGAACA TGAAAGGCAA AGATAAACGG    4625

ATGAACCCAT TACTAGCTTT GGCTGTATCA GACCAATAAC TTGAAATGCA CTTGTGCTAG    4685

CATGCCTAAG TATTAGAAAA GGTAGCATGG GAGAATCTAT ATTATTTTGG CTAACTTCTT    4745

TAGTTACTAT TGATTGATGA GAAAGCCTAC CATTGCCCAT GCCAGCCCTA ATGTCCCGGT    4805

GACATGATTG AGCCAGTACT ATGATTAATT TACTCTATTG TTCTCCTTTT TTGAGTGCTG    4865

TATAAGATGT CCTTTTTTTG AGCCACTCGA GAAGATGTTT ACTTAACTCT AGTGCGCAAT    4925

GATTGGAGCT CTCAGTGCAA CGCATGTGCT CTGTAATCTA CTGTCACCAC TACTCTGTAG    4985

TGTGTGCTTA AACTCTAAAC TATTCCACGT GGCTAGTAAT TACCAATCAT TTACAACACT    5045

GTTACATGTG TAG GGC TGC GAT GCT TCG GTG CTG TTG GAC AAC AGC AGC        5094
              Gly Cys Asp Ala Ser Val Leu Leu Asp Asn Ser Ser
                                  80                  85

AGC ATA GTT AGT GAG AAA GGG TCC AAC CCG AAC AGG AAC TCC CTC AGG      5142
Ser Ile Val Ser Glu Lys Gly Ser Asn Pro Asn Arg Asn Ser Leu Arg
        90                  95                  100

GGG TTT GAG GTG ATC GAC CAG ATT AAG GCT GCT CTT GAG GCT GCC TGC      5190
Gly Phe Glu Val Ile Asp Gln Ile Lys Ala Ala Leu Glu Ala Ala Cys
```

```
    105                 110                 115

CCA GGC ACA GTC TCC TGT GCC GAC ATT GTT GCC CTT GCG GCT CGT GAT      5238
Pro Gly Thr Val Ser Cys Ala Asp Ile Val Ala Leu Ala Ala Arg Asp
120                 125                 130                 135

TCC ACC GCC CTG GTATGTTCCA CTATCGACAA TCCTTTCCAA CCTCAAGGAA          5290
Ser Thr Ala Leu


CAGACATGAT ATTTGTGTGT GTGTGTGTGT GTATATATAT ATATAGTGAT AGCTTTGGCA    5350

AACTTAGATA TTTTCTGAGC TCTAAACCGT AG GTT GGT GGA CCA TAC TGG GAC      5403
                                    Val Gly Gly Pro Tyr Trp Asp
                                    140                 145

GTG CCA CTT GGC CGG AGA GAC TCG CTC GGT GCA AGC ATC CAG GGC TCC      5451
Val Pro Leu Gly Arg Arg Asp Ser Leu Gly Ala Ser Ile Gln Gly Ser
            150                 155                 160

AAC AAT GAC ATC CCA GCC CCC AAC AAC ACA CTC CCC ACT ATC ATC ACC      5499
Asn Asn Asp Ile Pro Ala Pro Asn Asn Thr Leu Pro Thr Ile Ile Thr
        165                 170                 175

AAG TTC AAG CGC CAG GGC CTC AAT GTT GTT GAT GTT GTC GCC CTC TCA      5547
Lys Phe Lys Arg Gln Gly Leu Asn Val Val Asp Val Val Ala Leu Ser
    180                 185                 190

GGTGATTTTT CTTGTATTTA TTAGTAACAT CTGTCCTTCG TTATTCACCA ACTTAGCGCA    5607

CACTCATATT ACGCATGGAT ACAATATCAT GTGTGAATAC A GGT GGT CAC ACC        5660
                                      Gly Gly His Thr
                                      195

ATT GGT ATG TCT CGG TGC ACT AGT TTC CGG CAG AGG CTA TAC AAC CAG      5708
Ile Gly Met Ser Arg Cys Thr Ser Phe Arg Gln Arg Leu Tyr Asn Gln
    200                 205                 210

ACA GGC AAT GGC ATG GCT GAC AGC ACA CTG GAT GTA TCC TAC GCC GCA      5756
Thr Gly Asn Gly Met Ala Asp Ser Thr Leu Asp Val Ser Tyr Ala Ala
215                 220                 225                 230

AAG CTG AGG CAG GGA TGC CCC CGC TCT GGT GGT GAC AAC AAC CTC TTC      5804
Lys Leu Arg Gln Gly Cys Pro Arg Ser Gly Gly Asp Asn Asn Leu Phe
                235                 240                 245

CCC TTG GAC TTC ATC ACC CCT GCC AAG TTT GAC AAT TTT TAC TAC AAG      5852
Pro Leu Asp Phe Ile Thr Pro Ala Lys Phe Asp Asn Phe Tyr Tyr Lys
            250                 255                 260

AAC CTC CTG GCC GGC AAG GGC CTT CTA AGC TCT GAT GAG ATT CTG TTA      5900
Asn Leu Leu Ala Gly Lys Gly Leu Leu Ser Ser Asp Glu Ile Leu Leu
        265                 270                 275

ACC AAG AGC GCT GAG ACA GCG GCC CTC GTG AAG GCA TAT GCT GCT GAT      5948
Thr Lys Ser Ala Glu Thr Ala Ala Leu Val Lys Ala Tyr Ala Ala Asp
    280                 285                 290
```

```
GTC AAT CTC TTC TTC CAG CAC TTT GCA CAG TCT ATG GTG AAT ATG GGA      5996
Val Asn Leu Phe Phe Gln His Phe Ala Gln Ser Met Val Asn Met Gly
295                 300                 305                 310

AAC ATC TCG CCA CTG ACA GGG TCA CAA GGT GAG ATC AGG AAG AAC TGC      6044
Asn Ile Ser Pro Leu Thr Gly Ser Gln Gly Glu Ile Arg Lys Asn Cys
                315                 320                 325

AGG AGG CTC AAC AAT GAC CAC TGA GGGCACTGAA GTCGCTTGAT GTGCTGAATT     6098
Arg Arg Leu Asn Asn Asp His  *
            330

GTTCGTGATG TTGGTGGCGT ATTTTGTTTA AATAAGTAAG CATGGCTGTG ATTTTATCAT    6158

ATGATCGATC TTTGGGGTTT TATTTAACAC ATTGTAAAAT GTGTATCTAT TAATAACTCA    6218

ATGTATAAGA TGTGTTCATT CTTCGGTTGC CATAGATCTG CTTATTTGAC CTGTGATGTT    6278

TTGACTCCAA AAACCAAAAT CACAACTCAA TAAACTCATG GAATATGTCC ACCTGTTTCT    6338

TGAAGAGTTC ATCTACCATT CCAGTTGGCA TTTATCAGTG TTGCAGCGGC GCTGTGCTTT    6398

GTAACATAAC AATTGTTCAC GGCATATATC CAAATCTAGA GGCCTACCAA AATGAGATAA    6458

CAAGCCAACT AATCTGCTGG GAAATAGGTA ACAAGTCTCT AACAAGATCC GTTGACCTGC    6518

AGGTCGACCT CGAGGGGGGG CCCGGTACCC AA                                  6550
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 334 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Ala Thr Ser Met Gly Cys Leu Val Leu Leu Cys Leu Val Ser Ser
  1               5                  10                  15

Leu Leu Pro Ser Ala Val Leu Gly His Pro Trp Gly Gly Leu Phe Pro
             20                  25                  30

Gln Phe Tyr Asp His Ser Cys Pro Lys Ala Lys Glu Ile Val Gln Ser
         35                  40                  45

Ile Val Ala Gln Ala Val Ala Lys Glu Thr Arg Met Ala Ala Ser Leu
     50                  55                  60

Val Arg Leu His Phe His Asp Cys Phe Val Lys Gly Cys Asp Ala Ser
 65                  70                  75                  80

Val Leu Leu Asp Asn Ser Ser Ser Ile Val Ser Glu Lys Gly Ser Asn
```

```
                85                  90                  95

Pro Asn Arg Asn Ser Leu Arg Gly Phe Glu Val Ile Asp Gln Ile Lys
            100                 105                 110

Ala Ala Leu Glu Ala Ala Cys Pro Gly Thr Val Ser Cys Ala Asp Ile
        115                 120                 125

Val Ala Leu Ala Ala Arg Asp Ser Thr Ala Leu Val Gly Gly Pro Tyr
    130                 135                 140

Trp Asp Val Pro Leu Gly Arg Arg Asp Ser Leu Gly Ala Ser Ile Gln
145                 150                 155                 160

Gly Ser Asn Asn Asp Ile Pro Ala Pro Asn Asn Thr Leu Pro Thr Ile
                165                 170                 175

Ile Thr Lys Phe Lys Arg Gln Gly Leu Asn Val Val Asp Val Val Ala
            180                 185                 190

Leu Ser Gly Gly His Thr Ile Gly Met Ser Arg Cys Thr Ser Phe Arg
        195                 200                 205

Gln Arg Leu Tyr Asn Gln Thr Gly Asn Gly Met Ala Asp Ser Thr Leu
    210                 215                 220

Asp Val Ser Tyr Ala Ala Lys Leu Arg Gln Gly Cys Pro Arg Ser Gly
225                 230                 235                 240

Gly Asp Asn Asn Leu Phe Pro Leu Asp Phe Ile Thr Pro Ala Lys Phe
                245                 250                 255

Asp Asn Phe Tyr Tyr Lys Asn Leu Leu Ala Gly Lys Gly Leu Leu Ser
            260                 265                 270

Ser Asp Glu Ile Leu Leu Thr Lys Ser Ala Glu Thr Ala Ala Leu Val
        275                 280                 285

Lys Ala Tyr Ala Ala Asp Val Asn Leu Phe Phe Gln His Phe Ala Gln
    290                 295                 300

Ser Met Val Asn Met Gly Asn Ile Ser Pro Leu Thr Gly Ser Gln Gly
305                 310                 315                 320

Glu Ile Arg Lys Asn Cys Arg Arg Leu Asn Asn Asp His  *
                325                 330
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TTYCAYGAYT GYTTYGTYAA YGGBTG 26

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

SGTRTGSGCS CCGSWSAGVG CSAC 24

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1354 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATCAACCAGC AACACTCTTC TCTTATAACA TAGTACAGCG AAGGTAACTC ACATGGCAAC 60

TTCCATGGGT TGTCTCGTCT TGCTCTGCCT TGTTTCTTCT CTCCTTCCCA GTGCCGTCCT 120

TGGCCACCCA TGGGGTGGCT TGTTCCCACA GTTCTATGAC CATTCGTGCC CCAAGGCGAA 180

GGAGATTGTG CAGTCCATTG TGGCACAGGC TGTGGCCAAG GAGACCAGGA TGGCGGCATC 240

TTTAGTCAGA CTGCATTTCC ATGACTGCTT TGTCAAGGGC TGCGATGCTT CGGTGCTGTT 300

GGACAACAGC AGCAGCATAG TTAGTGAGAA AGGGTCCAAC CCGAACAGGA ACTCCCTCAG 360

GGGGTTTGAG GTGATCGACC AGATTAAGGC TGCTCTTGAG GCTGCCTGCC CAGGCACAGT 420

CTCCTGTGCC GACATTGTTG CCCTTGCGGC TCGTGATTCC ACCGCCCTGG TTGGTGGACC 480

ATACTGGGAC GTGCCACTTG GCCGGAGAGA CTCGCTCGGT GCAAGCATCC AGGGCTCCAA 540

CAATGACATC CCAGCCCCCA ACAACACACT CCCCACTATC ATCACCAAGT TCAAGCGCCA 600

GGGCCTCAAT GTTGTTGATG TTGTCGCCCT CTCAGGTGGT CACACCATTG GTATGTCTCG 660

GTGCACTAGT TTCCGGCAGA GGCTATACAA CCAGACAGGC AATGGCATGG CTGACAGCAC 720

ACTGGATGTA TCCTACGCCG CAAAGCTGAG GCAGGGATGC CCCCGCTCTG GTGGTGACAA 780

CAACCTCTTC CCCTTGGACT TCATCACCCC TGCCAAGTTT GACAATTTTT ACTACAAGAA 840

CCTCCTGGCC GGCAAGGGCC TTCTAAGCTC TGATGAGATT CTGTTAACCA AGAGCGCTGA 900

GACAGCGGCC CTCGTGAAGG CATATGCTGC TGATGTCAAT CTCTTCTTCC AGCACTTTGC 960

ACAGTCTATG GTGAATATGG GAAACATCTC GCCACTGACA GGGTCACAAG GTGAGATCAG 1020

GAAGAACTGC AGGAGGCTCA ACAATGACCA CTGAGGGCAC TGAAGTCGCT TGATGTGCTG 1080

AATTGTTCGT GATGTTGGTG GCGTATTTTG TTTAAATAAG TAAGCATGGC TGTGATTTTA 1140

TCATATGATC GATCTTTGGG GTTTTATTTA ACACATTGTA AAATGTGTAT CTATTAATAA 1200

CTCAATGTAT AAGATGTGTT CATTCTTCGG TTGCCATAGA TCTGCTTATT TGACCTGTGA 1260

TGTTTTGACT CCAAAAACCA AAATCACAAC TCAATAAACT CATGGAATAT GTCCACCTGT 1320

TTCTTGAAAA AAAAAAAAAA AAAAAAAAAA AAAA 1354

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GTCATAGAAC TGTGGG 16

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATAACATAGT ACAGCG 16

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10160 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGGCCCGCTA GCGGTACCCC CGGGGTCGAC CATGGTCCGT CCTGTAGAAA CCCCAACCCG 60

TGAAATCAAA AAACTCGACG GCCTGTGGGC ATTCAGTCTG GATCGCGAAA ACTGTGGAAT 120

TGATCAGCGT TGGTGGGAAA GCGCGTTACA AGAAAGCCGG GCAATTGCTG TGCCAGGCAG 180

TTTTAACGAT CAGTTCGCCG ATGCAGATAT TCGTAATTAT GCGGGCAACG TCTGGTATCA 240

GCGCGAAGTC TTTATACCGA AAGGTTGGGC AGGCCAGCGT ATCGTGCTGC GTTTCGATGC 300

GGTCACTCAT TACGGCAAAG TGTGGGTCAA TAATCAGGAA GTGATGGAGC ATCAGGGCGG 360

CTATACGCCA TTTGAAGCCG ATGTCACGCC GTATGTTATT GCCGGGAAAA GTGTACGTAT 420

CACCGTTTGT GTGAACAACG AACTGAACTG GCAGACTATC CCGCCGGGAA TGGTGATTAC 480

CGACGAAAAC GGCAAGAAAA AGCAGTCTTA CTTCCATGAT TTCTTTAACT ATGCCGGAAT 540

CCATCGCAGC GTAATGCTCT ACACCACGCC GAACACCTGG GTGGACGATA TCACCGTGGT 600

GACGCATGTC GCGCAAGACT GTAACCACGC GTCTGTTGAC TGGCAGGTGG TGGCCAATGG 660

TGATGTCAGC GTTGAACTGC GTGATGCGGA TCAACAGGTG GTTGCAACTG GACAAGGCAC 720

TAGCGGGACT TTGCAAGTGG TGAATCCGCA CCTCTGGCAA CCGGGTGAAG GTTATCTCTA 780

TGAACTGTGC GTCACAGCCA AAAGCCAGAC AGAGTGTGAT ATCTACCCGC TTCGCGTCGG 840

CATCCGGTCA GTGGCAGTGA AGGGCGAACA GTTCCTGATT AACCACAAAC CGTTCTACTT 900

TACTGGCTTT GGTCGTCATG AAGATGCGGA CTTACGTGGC AAAGGATTCG ATAACGTGCT 960

GATGGTGCAC GACCACGCAT TAATGGACTG GATTGGGGCC AACTCCTACC GTACCTCGCA 1020

TTACCCTTAC GCTGAAGAGA TGCTCGACTG GGCAGATGAA CATGGCATCG TGGTGATTGA 1080

TGAAACTGCT GCTGTCGGCT TTAACCTCTC TTTAGGCATT GGTTTCGAAG CGGGCAACAA 1140

GCCGAAAGAA CTGTACAGCG AAGAGGCAGT CAACGGGGAA ACTCAGCAAG CGCACTTACA 1200

GGCGATTAAA GAGCTGATAG CGCGTGACAA AAACCACCCA AGCGTGGTGA TGTGGAGTAT 1260

TGCCAACGAA CCGGATACCC GTCCGCAAGT GCACGGGAAT ATTTCGCCAC TGGCGGAAGC 1320

AACGCGTAAA CTCGACCCGA CGCGTCCGAT CACCTGCGTC AATGTAATGT TCTGCGACGC 1380

TCACACCGAT ACCATCAGCG ATCTCTTTGA TGTGCTGTGC CTGAACCGTT ATTACGGATG 1440

GTATGTCCAA AGCGGCGATT TGGAAACGGC AGAGAAGGTA CTGGAAAAAG AACTTCTGGC 1500

CTGGCAGGAG AAACTGCATC AGCCGATTAT CATCACCGAA TACGGCGTGG ATACGTTAGC 1560

CGGGCTGCAC TCAATGTACA CCGACATGTG GAGTGAAGAG TATCAGTGTG CATGGCTGGA 1620

TATGTATCAC CGCGTCTTTG ATCGCGTCAG CGCCGTCGTC GGTGAACAGG TATGGAATTT 1680

CGCCGATTTT GCGACCTCGC AAGGCATATT GCGCGTTGGC GGTAACAAGA AAGGGATCTT 1740

CACTCGCGAC CGCAAACCGA AGTCGGCGGC TTTTCTGCTG CAAAAACGCT GGACTGGCAT 1800

GAACTTCGGT GAAAAACCGC AGCAGGGAGG CAAACAATGA ATCAACAACT CTCCTGGCGC 1860

ACCATCGTCG GCTACAGCCT CGGTGGGGAA TTGGAGCTCG AATTTCCCCG ATCGTTCAAA 1920

CATTTGGCAA TAAAGTTTCT TAAGATTGAA TCCTGTTGCC GGTCTTGCGA TGATTATCAT 1980

ATAATTTCTG TTGAATTACG TTAAGCATGT AATAATTAAC ATGTAATGCA TGACGTTATT 2040

TATGAGATGG GTTTTTATGA TTAGAGTCCC GCAATTATAC ATTTAATACG CGATAGAAAA 2100

CAAAATATAG CGCGCAAACT AGGATAAATT ATCGCGCGCG GTGTCATCTA TGTTACTAGA 2160

TCGATCGGGA ATTAAGCTTA GATCTGCATG GGTGGAGACT TTTCAACAAA GGGTAATATC 2220

CGGAAACCTC CTCGGATTCC ATTGCCCAGC TATCTGTCAC TTTATTGTGA AGATAGTGGA 2280

AAAGGAAGGT GGCTCCTACA AATGCCATCA TTGCGATAAA GGAAAGGCCA TCGTTGAAGA 2340

TGCCTCTGCC GACAGTGGTC CCAAAGATGG ACCCCCACCC ACGAGGAGCA TCGTGGAAAA 2400

AGAAGACGTT CCAACCACGT CTTCAAAGCA AGTGGATTGA TGTGATCATC GATGGAGACT 2460

TTTCAACAAA GGGTAATATC CGGAAACCTC CTCGGATTCC ATTGCCCAGC TATCTGTCAC 2520

TTTATTGTGA AGATAGTGGA AAAGGAAGGT GGCTCCTACA AATGCCATCA TTGCGATAAA 2580

GGAAAGGCCA TCGTTGAAGA TGCCTCTGCC GACAGTGGTC CCAAAGATGG ACCCCCACCC 2640

ACGAGGAGCA TCGTGGAAAA AGAAGACGTT CCAACCACGT CTTCAAAGCA AGTGGATTGA 2700

TGTGATATCT CCACTGACGT AAGGGATGAC GCACAATCCC ACTATCCTTC GCAAGACCCT 2760

TCCTCTATAT AAGGAAGTTC ATTTCATTTG GAGAGAACAC GGGGGACTCT AGAGGATCCA 2820

GCTGAAGGCT CGACAAGGCA GTCCACGGAG GAGCTGATAT TTGGTGGACA AGCTGTGGAT 2880

AGGAGCAACC CTATCCCTAA TATACCAGCA CCACCAAGTC AGGGCAATCC CCAGATCAAG 2940

TGCAAAGGTC CGCCTTGTTT CTCCTCTGTC TCTTGATCTG ACTAATCTTG GTTTATGATT 3000

CGTTGAGTAA TTTTGGGGAA AGCTCCTTTG CTGCTCCACA CATGTCCATT CGAATTTTAC 3060

CGTGTTTAGC AAGGGCGAAA AGTTTGCATC TTGATGATTT AGCTTGACTA TGCGATTGCT 3120

TTCCTGGACC CGTGCAGCTG CGCTCGGATC TGGGGCCATT TGTTCCAGGC ACGGGATAAG 3180

CATTCAGCCA TGGCAGACGC CAAAAACATA AAGAAAGGCC CGGCGCCATT CTATCCTCTA 3240

GAGGATGGAA CCGCTGGAGA GCAACTGCAT AAGGCTATGA AGAGATACGC CCTGGTTCCT 3300

GGAACAATTG CTTTTACAGA TGCACATATC GAGGTGAACA TCACGTACGC GGAATACTTC 3360

GAAATGTCCG TTCGGTTGGC AGAAGCTATG AAACGATATG GGCTGAATAC AAATCACAGA 3420

ATCGTCGTAT GCAGTGAAAA CTCTCTTCAA TTCTTTATGC CGGTGTTGGG CGCGTTATTT 3480

ATCGGAGTTG CAGTTGCGCC CGCGAACGAC ATTTATAATG AACGTGAATT GCTCAACAGT 3540

ATGAACATTT CGCAGCCTAC CGTAGTGTTT GTTTCCAAAA AGGGGTTGCA AAAAATTTTG 3600

AACGTGCAAA AAAAATTACC AATAATCCAG AAAATTATTA TCATGGATTC TAAAACGGAT 3660

TACCAGGGAT TTCAGTCGAT GTACACGTTC GTCACATCTC ATCTACCTCC CGGTTTTAAT 3720

GAATACGATT TTGTACCAGA GTCCTTTGAT CGTGACAAAA CAATTGCACT GATAATGAAT 3780

TCCTCTGGAT CTACTGGGTT ACCTAAGGGT GTGGCCCTTC CGCATAGAAC TGCCTGCGTC 3840

AGATTCTCGC ATGCCAGAGA TCCTATTTTT GGCAATCAAA TCATTCCGGA TACTGCGATT 3900

TTAAGTGTTG TTCCATTCCA TCACGGTTTT GGAATGTTTA CTACACTCGG ATATTTGATA 3960

TGTGGATTTC GAGTCGTCTT AATGTATAGA TTTGAAGAAG AGCTGTTTTT ACGATCCCTT 4020

CAGGATTACA AAATTCAAAG TGCGTTGCTA GTACCAACCC TATTTTCATT CTTCGCCAAA 4080

AGCACTCTGA TTGACAAATA CGATTTATCT AATTTACACG AAATTGCTTC TGGGGGCGCA 4140

CCTCTTTCGA AAGAAGTCGG GGAAGCGGTT GCAAAACGCT TCCATCTTCC AGGGATACGA 4200

CAAGGATATG GGCTCACTGA GACTACATCA GCTATTCTGA TTACACCCGA GGGGGATGAT 4260

AAACCGGGCG CGGTCGGTAA AGTTGTTCCA TTTTTTGAAG CGAAGGTTGT GGATCTGGAT 4320

ACCGGGAAAA CGCTGGGCGT TAATCAGAGA GGCGAATTAT GTGTCAGAGG ACCTATGATT 4380

ATGTCCGGTT ATGTAAACAA TCCGGAAGCG ACCAACGCCT TGATTGACAA GGATGGATGG 4440

CTACATTCTG GAGACATAGC TTACTGGGAC GAAGACGAAC ACTTCTTCAT AGTTGACCGC 4500

TTGAAGTCTT TAATTAAATA CAAAGGATAT CAGGTGGCCC CCGCTGAATT GGAATCGATA 4560

TTGTTACAAC ACCCCAACAT CTTCGACGCG GGCGTGGCAG GTCTTCCCGA CGATGACGCC 4620

GGTGAACTTC CCGCCGCCGT TGTTGTTTTG GAGCACGGAA AGACGATGAC GGAAAAAGAG 4680

ATCGTGGATT ACGTCGCCAG TCAAGTAACA ACCGCGAAAA AGTTGCGCGG AGGAGTTGTG 4740

TTTGTGGACG AAGTACCGAA AGGTCTTACC GGAAAACTCG ACGCAAGAAA AATCAGAGAG 4800

ATCCTCATAA AGGCCAAGAA GGGCGGAAAG TCCAAATTGT AAAATGTAAC TGTATTCAGC 4860

GATGACGAAA TTCTTAGCTA TTGTAATCAG ATCCGCGAAT TTCCCCGATC GTTCAAACAT 4920

TTGGCAATAA AGTTTCTTAA GATTGAATCC TGTTGCCGGT CTTGCGATGA TTATCATATA 4980

ATTTCTGTTG AATTACGTTA AGCATGTAAT AATTAACATG TAATGCATGA CGTTATTTAT 5040

GAGATGGGTT TTTATGATTA GAGTCCCGCA ATTATACATT TAATACGCGA TAGAAAACAA 5100

AATATAGCGC GCAAACTAGG ATAAATTATC GCGCGCGGTG TCATCTATGT TACTAGATCG 5160

ATCGGGAATT GAGATCTCAT ATGTCGAGCT CGGGGATCTC CTTTGCCCCA GAGATCACAA 5220

TGGACGACTT CCTCTATCTC TACGATCTAG TCAGGAAGTT CGACGGAGAA GGTGACGATA 5280

CCATGTTCAC CACTGATAAT GAGAAGATTA GCCTTTTCAA TTTCAGAAAG AATGCTAACC 5340

CACAGATGGT TAGAGAGGCT TACGCAGCAG GTCTCATCAA GACGATCTAC CCGAGCAATA 5400

ATCTCCAGGA GATCAAATAC CTTCCCAAGA AGGTTAAAGA TGCAGTCAAA AGATTCAGGA 5460

CTAACTGCAT CAAGAACACA GAGAAAGATA TATTTCTCAA GATCAGAAGT ACTATTCCAG 5520

TATGGACGAT TCAAGGCTTG CTTCACAAAC CAAGGCAAGT AATAGAGATT GGAGTCTCTA 5580

AAAAGGTAGT TCCCACTGAA TCAAAGGCCA TGGAGTCAAA GATTCAAATA GAGGACCTAA 5640

CAGAACTCGC CGTAAAGACT GGCGAACAGT TCCATCGATG ATTGAGACTT TTCAACAAAG 5700

GGTAATATCC GGAAACCTCC TCGGATTCCA TTGCCCAGCT ATCTGTCACT TTATTGTGAA 5760

GATAGTGGAA AAGGAAGGTG GCTCCTACAA ATGCCATCAT TGCGATAAAG GAAAGGCCAT 5820

CGTTGAAGAT GCCTCTGCCG ACAGTGGTCC CAAAGATGGA CCCCCACCCA CGAGGAGCAT 5880

CGTGGAAAAA GAAGACGTTC CAACCACGTC TTCAAAGCAA GTGGATTGAT GTGATATCTC 5940

CACTGACGTA AGGGATGACG CACAATCCCA CTATCCTTCG CAAGACCCTT CCTCTATATA 6000

AGGAAGTTCA TTTCATTTGG AGAGGACACG CTGACAAGCT CGGATCCTTT AGCATGATTG 6060

AACAAGATGG ATTGCACGCA GGTTCTCCGG CCGCTTGGGT GGAGAGGCTA TTCGGCTATG 6120

ACTGGGCACA ACAGACAATC GGCTGCTCTG ATGCCGCCGT GTTCCGGCTG TCAGCGCAGG 6180

GGCGCCCGGT TCTTTTTGTC AAGACCGACC TGTCCGGTGC CCTGAATGAA CTGCAGGACG 6240

AGGCAGCGCG GCTATCGTGG CTGGCCACGA CGGGCGTTCC TTGCGCAGCT GTGCTCGACG 6300

TTGTCACTGA AGCGGGAAGG GACTGGCTGC TATTGGGCGA AGTGCCGGGG CAGGATCTCC 6360

TGTCATCTCA CCTTGCTCCT GCCGAGAAAG TATCCATCAT GGCTGATGCA ATGCGGCGGC 6420

TGCATACGCT TGATCCGGCT ACCTGCCCAT TCGACCACCA AGCGAAACAT CGCATCGAGC 6480

GAGCACGTAC TCGGATGGAA GCCGGTCTTG TCGATCAGGA TGATCTGGAC GAAGAGCATC 6540

AGGGGCTCGC GCCAGCCGAA CTGTTCGCCA GGCTCAAGGC GCGCATGCCC GACGGCGAGG 6600

ATCTCGTCGT GACCCATGGC GATGCCTGCT TGCCGAATAT CATGGTGGAA AATGGCCGCT 6660

TTTCTGGATT CATCGACTGT GGCCGGCTGG GTGTGGCGGA CCGCTATCAG GACATAGCGT 6720

TGGCTACCCG TGATATTGCT GAAGAGCTTG GCGGCGAATG GGCTGACCGC TTCCTCGTGC 6780

TTTACGGTAT CGCCGCTCCC GATTCGCAGC GCATCGCCTT CTATCGCCTT CTTGACGAGT 6840

TCTTCTGAGC GGGACTCTGG GGTTCGAAAT GACCGACCAA GCGACGCCCA ACCTGCCATC 6900

ACGAGATTTC GATTCCACCG CCGCCTTCTA TGAAAGGTTG GGCTTCGGAA TCGTTTTCCG 6960

GGACGCCGGC TGGATGATCC TCCAGCGCGG GGATCTCATG CTGGAGTTCT TCGCCCACCC 7020

CAACAGAGGT GGATGGACAG ACCCGTTCTT ACACCGGACT GGGCGCGGGA TAGGATATTC 7080

AGATTGGGAT GGGATTGAGC TTAAAGCCGG CGCTGAGACC ATGCTCAAGG TAGGCAATGT 7140

CCTCAGCGTC GAGCCCGGCA TCTATGTCGA GGGCATTGGT GGAGCGCGCT TCGGGGATAC 7200

CGTGCTTGTA ACTGAGACCG GATATGAGGC CCTCACTCCG CTTGATCTTG GCAAAGATAT 7260

TTGACGCATT TATTAGTATG TGTTAATTTT CATTTGCAGT GCAGTATTTT CTATTCGATC 7320

TTTATGTAAT TCGTTACAAT TAATAAATAT TCAAATCAGA TTATTGACTG TCATTTGTAT 7380

CAAATCGTGT TTAATGGATA TTTTTATTAT AATATTGATG ATATCTCAAT CAAAACGTAG 7440

ATAATAATAA TATTTATTTA ATATTTTTGC GTCGCACAGT GAAAATCTAT ATGAGATTAC 7500

AAAATACCGA CAACATTATT TAAGATACAT AGACATTAAC CCTGAGACTG TTGGACATCA 7560

ACGGGTAGAT TCCTTCATGC ATAGCACCTC ATTCTTGGGG ACAAAAGCAC GGTTTGGCCG 7620

TTCCATTGCT GCACGAACGA GCTTTGCTAT ATCCTCGGGT TGGATCATCT CATCAGGTCC 7680

AATCAAATTT GTCCAAGAAC TCATGTTAGT CGCAACGAAA CCGGGGCATA TGGTGCACTC 7740

TCAGTACAAT CTGCTCTGAT GCCGCATAGT TAAGCCAGCC CCGACACCCG CCAACACCCG 7800

CTGACGCGCC CTGACGGGCT TGTCTGCTCC CGGCATCCGC TTACAGACAA GCTGTGACCG 7860

TCTCCGGGAG CTGCATGTGT CAGAGGTTTT CACCGTCATC ACCGAAACGC GCGAGACGAA 7920

AGGGCCTCGT GATACGCCTA TTTTTATAGG TTAATGTCAT GATAATAATG GTTTCTTAGA 7980

CGTCAGGTGG CACTTTTCGG GGAAATGTGC GCGGAACCCC TATTTGTTTA TTTTTCTAAA 8040

TACATTCAAA TATGTATCCG CTCATGAGAC AATAACCCTG ATAAATGCTT CAATAATATT 8100

GAAAAAGGAA GAGTATGAGT ATTCAACATT TCCGTGTCGC CCTTATTCCC TTTTTTGCGG 8160

CATTTTGCCT TCCTGTTTTT GCTCACCCAG AAACGCTGGT GAAAGTAAAA GATGCTGAAG 8220

ATCAGTTGGG TGCACGAGTG GGTTACATCG AACTGGATCT CAACAGCGGT AAGATCCTTG 8280

AGAGTTTTCG CCCCGAAGAA CGTTTTCCAA TGATGAGCAC TTTTAAAGTT CTGCTATGTG 8340

GCGCGGTATT ATCCCGTATT GACGCCGGGC AAGAGCAACT CGGTCGCCGC ATACACTATT 8400

CTCAGAATGA CTTGGTTGAG TACTCACCAG TCACAGAAAA GCATCTTACG GATGGCATGA 8460

CAGTAAGAGA ATTATGCAGT GCTGCCATAA CCATGAGTGA TAACACTGCG GCCAACTTAC 8520

TTCTGACAAC GATCGGAGGA CCGAAGGAGC TAACCGCTTT TTTGCACAAC ATGGGGGATC 8580

ATGTAACTCG CCTTGATCGT TGGGAACCGG AGCTGAATGA AGCCATACCA AACGACGAGC 8640

GTGACACCAC GATGCCTGTA GCAATGGCAA CAACGTTGCG CAAACTATTA ACTGGCGAAC 8700

TACTTACTCT AGCTTCCCGG CAACAATTAA TAGACTGGAT GGAGGCGGAT AAAGTTGCAG 8760

GACCACTTCT GCGCTCGGCC CTTCCGGCTG GCTGGTTTAT TGCTGATAAA TCTGGAGCCG 8820

GTGAGCGTGG GTCTCGCGGT ATCATTGCAG CACTGGGGCC AGATGGTAAG CCCTCCCGTA 8880

TCGTAGTTAT CTACACGACG GGGAGTCAGG CAACTATGGA TGAACGAAAT AGACAGATCG 8940

CTGAGATAGG TGCCTCACTG ATTAAGCATT GGTAACTGTC AGACCAAGTT TACTCATATA 9000

TACTTTAGAT TGATTTAAAA CTTCATTTTT AATTTAAAAG GATCTAGGTG AAGATCCTTT 9060

TTGATAATCT CATGACCAAA ATCCCTTAAC GTGAGTTTTC GTTCCACTGA GCGTCAGACC 9120

CCGTAGAAAA GATCAAAGGA TCTTCTTGAG ATCCTTTTTT TCTGCGCGTA ATCTGCTGCT 9180

TGCAAACAAA AAAACCACCG CTACCAGCGG TGGTTTGTTT GCCGGATCAA GAGCTACCAA 9240

CTCTTTTTCC GAAGGTAACT GGCTTCAGCA GAGCGCAGAT ACCAAATACT GTCCTTCTAG 9300

TGTAGCCGTA GTTAGGCCAC CACTTCAAGA ACTCTGTAGC ACCGCCTACA TACCTCGCTC 9360

TGCTAATCCT GTTACCAGTG GCTGCTGCCA GTGGCGATAA GTCGTGTCTT ACCGGGTTGG 9420

ACTCAAGACG ATAGTTACCG GATAAGGCGC AGCGGTCGGG CTGAACGGGG GGTTCGTGCA 9480

CACAGCCCAG CTTGGAGCGA ACGACCTACA CCGAACTGAG ATACCTACAG CGTGAGCATT 9540

GAGAAAGCGC CACGCTTCCC GAAGGGAGAA AGGCGGACAG GTATCCGGTA AGCGGCAGGG 9600

TCGGAACAGG AGAGCGCACG AGGGAGCTTC CAGGGGGAAA CGCCTGGTAT CTTTATAGTC 9660

CTGTCGGGTT TCGCCACCTC TGACTTGAGC GTCGATTTTT GTGATGCTCG TCAGGGGGGC 9720

GGAGCCTATG GAAAAACGCC AGCAACGCGG CCTTTTTACG GTTCCTGGCC TTTTGCTGGC 9780

CTTTTGCTCA CATGTTCTTT CCTGCGTTAT CCCCTGATTC TGTGGATAAC CGTATTACCG 9840

CCTTTGAGTG AGCTGATACC GCTCGCCGCA GCCGAACGAC CGAGCGCAGC GAGTCAGTGA 9900

GCGAGGAAGC GGAAGAGCGC CCAATACGCA AACCGCCTCT CCCCGCGCGT TGGCCGATTC 9960

ATTAATGCAG CTGGCACGAC AGGTTTCCCG ACTGGAAAGC GGGCAGTGAG CGCAACGCAA 10020

TTAATGTGAG TTAGCTCACT CATTAGGCAC CCCAGGCTTT ACACTTTATG CTTCCGGCTC 10080

GTATGTTGTG TGGAATTGTG AGCGGATAAC AATTTCACAC AGGAAACAGC TATGACCATG 10140

ATTACGCCAA GCTTCCGCGG 10160

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11784 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGGCCCACCA CTGTTGTAAC TTGTAAGCCA CTAGCTCACG TTCTCCATGA GCTCTTCTCT 60

CTGCTGTTTC TTCCTCTGCT AACTGCGTTA TGATATGACG TCGTATAAAT AATCTCACAA 120

TACTTCCTTA TTTTCAGCAT GGCCTCTTTT ATGTTTATTT AACAGTAGCA ACCAACGCCG 180

CTCGATGTTT CCTTCAAGAA ACGGCCACTC ACTATGTGGT GTGCAGAAGA ACAAATGTAA 240

GCAGCTCCTA CAGGTACCAG TAGTCATGTC AGTGTGGAAG CTTTCCAACC AACGCCTCCT 300

TCGAGGAACC TGGTCGTGCT GACATGAATG TAGGCCATGC AAGCACAAGC ACCTAACGCG 360

AATCATCACG ACGCGCCGTG TACTGGGCGT TGGTACATCA CACCCCGCGT TTGACCTGAT 420

CGGAAGCATG CGTGTGTGTT GGCTGCAGGA CCGGCTATAG GTTTCCTGCA TTGGACAGCA 480

GAAGCCAGTC ATGTTAGGCA CTCACGCGCT CCTGCCGTTT GATGAATCAT CCGGTCTTTC 540

GTATTGATCA CTAGTTCACT ACGCTGATAT AGCAAATTTT AAGATGTGAA ACCACGAGAC 600

GAGCGATAAA TCTTAGACGT TACCTATCCA TATGAAGCTT GTGCGAAAAA AAGGCGTGCC 660

GCTGTAGCAT CATTCGTATA CACTTTTGTC CCCAAAGACA GGGATACGAA TCCATGCTCG 720

ACAGAACCCT CCCTTCCCTG CAGATAACGA CACTTAAGTA TAACAAAAGT AGTTGGATTA 780

TTTCAGAAGC AAAATCTCAC TTTTCGCTGG CCTTTTTGTA CTTTGGTTAC TTGAGTTCAG 840

ACAGTGTATG CTATATTGTC ATGTGCTGCG TAAGGTTTAA ATATGGTTCG ACAAATATAT 900

CAGTATATCA CTACTTTGTT ATGGGTGGGG CCTAGCACAA ACTTGATACA GCTAGGATAA 960

AGTTAGAACG ATGACTGATC TACTGTAAAG CGACACCTGT CCTGTTATGG TAGTTTAAGT 1020

CCATTCCTGG ACGACTCCAG ATCCAGGATA TGATGCTGTT ACATAATGCG ATTGTTCACA 1080

ATAAAATTGC ATGATGTTCT TCTACTCTTT AGGCAGTTTT GTTCAACAGG CAAGTTGCAT 1140

AATGCATGTG CATATATGAG CAGCATAATC ATCAATTAAT CATAGGTTCG TCATTTTAGT 1200

TTCACTCCTT CACATTATTC CAGCCCTTGA AGAAAAATGT AGCAGTGCTT GCTGTTTAAT 1260

AAGTGGCAGA GCTGTTTTCA CTCCACCTAC GCTTGTCTAG GACCAAAATT TTAATCTGTC 1320

ACTTTGAGCT AAAACTGAAG CACCAAACCG CTACAAAAGA ACGTAGGAGC TGAATTGTAA 1380

CTTGATGGGA TTACTATAGC AGTTGCTACA GTTCTAGCTA GCTACCTTAT TCTATACGCA 1440

TCACCCTAAC AACCCGGCTG ACTGCTGCAT CTGACCCCAC CGTCCCCTGC TCCAAACCAA 1500

CTCTCCTTTC CTTGCATGCA CTACACCCAC TTCCTGCAGC TATATATACC ACCATATGCC 1560

CATCTTATGA AACCATCCAC AAGAGGAGAA GAAACAATCA ACCAGCAACA CTCTTCTCTT 1620

ATAACATAGT ACAGCGAAGG TAACTCACGT CGACCATGGT CCGTCCTGTA GAAACCCCAA 1680

CCCGTGAAAT CAAAAAACTC GACGGCCTGT GGGCATTCAG TCTGGATCGC GAAAACTGTG 1740

GAATTGATCA GCGTTGGTGG GAAAGCGCGT TACAAGAAAG CCGGGCAATT GCTGTGCCAG 1800

GCAGTTTTAA CGATCAGTTC GCCGATGCAG ATATTCGTAA TTATGCGGGC AACGTCTGGT 1860

ATCAGCGCGA AGTCTTTATA CCGAAAGGTT GGGCAGGCCA GCGTATCGTG CTGCGTTTCG 1920

ATGCGGTCAC TCATTACGGC AAAGTGTGGG TCAATAATCA GGAAGTGATG GAGCATCAGG 1980

GCGGCTATAC GCCATTTGAA GCCGATGTCA CGCCGTATGT TATTGCCGGG AAAAGTGTAC 2040

GTATCACCGT TTGTGTGAAC AACGAACTGA ACTGGCAGAC TATCCCGCCG GGAATGGTGA 2100

TTACCGACGA AAACGGCAAG AAAAAGCAGT CTTACTTCCA TGATTTCTTT AACTATGCCG 2160

GAATCCATCG CAGCGTAATG CTCTACACCA CGCCGAACAC CTGGGTGGAC GATATCACCG 2220

TGGTGACGCA TGTCGCGCAA GACTGTAACC ACGCGTCTGT TGACTGGCAG GTGGTGGCCA 2280

ATGGTGATGT CAGCGTTGAA CTGCGTGATG CGGATCAACA GGTGGTTGCA ACTGGACAAG 2340

GCACTAGCGG GACTTTGCAA GTGGTGAATC CGCACCTCTG GCAACCGGGT GAAGGTTATC 2400

TCTATGAACT GTGCGTCACA GCCAAAAGCC AGACAGAGTG TGATATCTAC CCGCTTCGCG 2460

TCGGCATCCG GTCAGTGGCA GTGAAGGGCG AACAGTTCCT GATTAACCAC AAACCGTTCT 2520

ACTTTACTGG CTTTGGTCGT CATGAAGATG CGGACTTACG TGGCAAAGGA TTCGATAACG 2580

TGCTGATGGT GCACGACCAC GCATTAATGG ACTGGATTGG GGCCAACTCC TACCGTACCT 2640

CGCATTACCC TTACGCTGAA GAGATGCTCG ACTGGGCAGA TGAACATGGC ATCGTGGTGA 2700

TTGATGAAAC TGCTGCTGTC GGCTTTAACC TCTCTTTAGG CATTGGTTTC GAAGCGGGCA 2760

ACAAGCCGAA AGAACTGTAC AGCGAAGAGG CAGTCAACGG GGAAACTCAG CAAGCGCACT 2820

TACAGGCGAT TAAAGAGCTG ATAGCGCGTG ACAAAAACCA CCCAAGCGTG GTGATGTGGA 2880

GTATTGCCAA CGAACCGGAT ACCCGTCCGC AAGTGCACGG GAATATTTCG CCACTGGCGG 2940

AAGCAACGCG TAAACTCGAC CCGACGCGTC CGATCACCTG CGTCAATGTA ATGTTCTGCG 3000

ACGCTCACAC CGATACCATC AGCGATCTCT TTGATGTGCT GTGCCTGAAC CGTTATTACG 3060

GATGGTATGT CCAAAGCGGC GATTTGGAAA CGGCAGAGAA GGTACTGGAA AAAGAACTTC 3120

TGGCCTGGCA GGAGAAACTG CATCAGCCGA TTATCATCAC CGAATACGGC GTGGATACGT 3180

TAGCCGGGCT GCACTCAATG TACACCGACA TGTGGAGTGA AGAGTATCAG TGTGCATGGC 3240

TGGATATGTA TCACCGCGTC TTTGATCGCG TCAGCGCCGT CGTCGGTGAA CAGGTATGGA 3300

ATTTCGCCGA TTTTGCGACC TCGCAAGGCA TATTGCGCGT TGGCGGTAAC AAGAAAGGGA 3360

TCTTCACTCG CGACCGCAAA CCGAAGTCGG CGGCTTTTCT GCTGCAAAAA CGCTGGACTG 3420

GCATGAACTT CGGTGAAAAA CCGCAGCAGG GAGGCAAACA ATGAATCAAC AACTCTCCTG 3480

GCGCACCATC GTCGGCTACA GCCTCGGTGG GGAATTGGAG CTCGAATTTC CCCGATCGTT 3540

CAAACATTTG GCAATAAAGT TTCTTAAGAT TGAATCCTGT TGCCGGTCTT GCGATGATTA 3600

TCATATAATT TCTGTTGAAT TACGTTAAGC ATGTAATAAT TAACATGTAA TGCATGACGT 3660

TATTTATGAG ATGGGTTTTT ATGATTAGAG TCCCGCAATT ATACATTTAA TACGCGATAG 3720

AAAACAAAAT ATAGCGCGCA AACTAGGATA AATTATCGCG CGCGGTGTCA TCTATGTTAC 3780

TAGATCGATC GGGAATTAAG CTTAGATCTG CATGGGTGGA GACTTTTCAA CAAAGGGTAA 3840

TATCCGGAAA CCTCCTCGGA TTCCATTGCC CAGCTATCTG TCACTTTATT GTGAAGATAG 3900

TGGAAAAGGA AGGTGGCTCC TACAAATGCC ATCATTGCGA TAAAGGAAAG GCCATCGTTG 3960

AAGATGCCTC TGCCGACAGT GGTCCCAAAG ATGGACCCCC ACCCACGAGG AGCATCGTGG 4020

AAAAAGAAGA CGTTCCAACC ACGTCTTCAA AGCAAGTGGA TTGATGTGAT CATCGATGGA 4080

GACTTTTCAA CAAAGGGTAA TATCCGGAAA CCTCCTCGGA TTCCATTGCC CAGCTATCTG 4140

TCACTTTATT GTGAAGATAG TGGAAAAGGA AGGTGGCTCC TACAAATGCC ATCATTGCGA 4200

TAAAGGAAAG GCCATCGTTG AAGATGCCTC TGCCGACAGT GGTCCCAAAG ATGGACCCCC 4260

ACCCACGAGG AGCATCGTGG AAAAAGAAGA CGTTCCAACC ACGTCTTCAA AGCAAGTGGA 4320

TTGATGTGAT ATCTCCACTG ACGTAAGGGA TGACGCACAA TCCCACTATC CTTCGCAAGA 4380

CCCTTCCTCT ATATAAGGAA GTTCATTTCA TTTGGAGAGA ACACGGGGGA CTCTAGAGGA 4440

TCCAGCTGAA GGCTCGACAA GGCAGTCCAC GGAGGAGCTG ATATTTGGTG GACAAGCTGT 4500

GGATAGGAGC AACCCTATCC CTAATATACC AGCACCACCA AGTCAGGGCA ATCCCCAGAT 4560

CAAGTGCAAA GGTCCGCCTT GTTTCTCCTC TGTCTCTTGA TCTGACTAAT CTTGGTTTAT 4620

GATTCGTTGA GTAATTTTGG GGAAAGCTCC TTTGCTGCTC CACACATGTC CATTCGAATT 4680

TTACCGTGTT TAGCAAGGGC GAAAAGTTTG CATCTTGATG ATTTAGCTTG ACTATGCGAT 4740

TGCTTTCCTG GACCCGTGCA GCTGCGCTCG GATCTGGGGC CATTTGTTCC AGGCACGGGA 4800

TAAGCATTCA GCCATGGCAG ACGCCAAAAA CATAAAGAAA GGCCCGGCGC CATTCTATCC 4860

TCTAGAGGAT GGAACCGCTG GAGAGCAACT GCATAAGGCT ATGAAGAGAT ACGCCCTGGT 4920

TCCTGGAACA ATTGCTTTTA CAGATGCACA TATCGAGGTG AACATCACGT ACGCGGAATA 4980

CTTCGAAATG TCCGTTCGGT TGGCAGAAGC TATGAAACGA TATGGGCTGA ATACAAATCA 5040

CAGAATCGTC GTATGCAGTG AAAACTCTCT TCAATTCTTT ATGCCGGTGT TGGGCGCGTT 5100

ATTTATCGGA GTTGCAGTTG CGCCCGCGAA CGACATTTAT AATGAACGTG AATTGCTCAA 5160

CAGTATGAAC ATTTCGCAGC CTACCGTAGT GTTTGTTTCC AAAAAGGGGT TGCAAAAAAT 5220

TTTGAACGTG CAAAAAAAAT TACCAATAAT CCAGAAAATT ATTATCATGG ATTCTAAAAC 5280

GGATTACCAG GGATTTCAGT CGATGTACAC GTTCGTCACA TCTCATCTAC CTCCCGGTTT 5340

TAATGAATAC GATTTTGTAC CAGAGTCCTT TGATCGTGAC AAAACAATTG CACTGATAAT 5400

GAATTCCTCT GGATCTACTG GGTTACCTAA GGGTGTGGCC CTTCCGCATA GAACTGCCTG 5460

CGTCAGATTC TCGCATGCCA GAGATCCTAT TTTTGGCAAT CAAATCATTC CGGATACTGC 5520

GATTTTAAGT GTTGTTCCAT TCCATCACGG TTTTGGAATG TTTACTACAC TCGGATATTT 5580

GATATGTGGA TTTCGAGTCG TCTTAATGTA TAGATTTGAA GAAGAGCTGT TTTTACGATC 5640

CCTTCAGGAT TACAAAATTC AAAGTGCGTT GCTAGTACCA ACCCTATTTT CATTCTTCGC 5700

CAAAAGCACT CTGATTGACA AATACGATTT ATCTAATTTA CACGAAATTG CTTCTGGGGG 5760

CGCACCTCTT TCGAAAGAAG TCGGGGAAGC GGTTGCAAAA CGCTTCCATC TTCCAGGGAT 5820

ACGACAAGGA TATGGGCTCA CTGAGACTAC ATCAGCTATT CTGATTACAC CCGAGGGGGA 5880

TGATAAACCG GGCGCGGTCG GTAAAGTTGT TCCATTTTTT GAAGCGAAGG TTGTGGATCT 5940

GGATACCGGG AAAACGCTGG GCGTTAATCA GAGAGGCGAA TTATGTGTCA GAGGACCTAT 6000

GATTATGTCC GGTTATGTAA ACAATCCGGA AGCGACCAAC GCCTTGATTG ACAAGGATGG 6060

ATGGCTACAT TCTGGAGACA TAGCTTACTG GGACGAAGAC GAACACTTCT TCATAGTTGA 6120

CCGCTTGAAG TCTTTAATTA AATACAAAGG ATATCAGGTG GCCCCCGCTG AATTGGAATC 6180

GATATTGTTA CAACACCCCA ACATCTTCGA CGCGGGCGTG GCAGGTCTTC CCGACGATGA 6240

CGCCGGTGAA CTTCCCGCCG CCGTTGTTGT TTTGGAGCAC GGAAAGACGA TGACGGAAAA 6300

AGAGATCGTG GATTACGTCG CCAGTCAAGT AACAACCGCG AAAAAGTTGC GCGGAGGAGT 6360

TGTGTTTGTG GACGAAGTAC CGAAAGGTCT TACCGGAAAA CTCGACGCAA GAAAAATCAG 6420

AGAGATCCTC ATAAAGGCCA AGAAGGGCGG AAAGTCCAAA TTGTAAAATG TAACTGTATT 6480

CAGCGATGAC GAAATTCTTA GCTATTGTAA TCAGATCCGC GAATTTCCCC GATCGTTCAA 6540

ACATTTGGCA ATAAAGTTTC TTAAGATTGA ATCCTGTTGC CGGTCTTGCG ATGATTATCA 6600

TATAATTTCT GTTGAATTAC GTTAAGCATG TAATAATTAA CATGTAATGC ATGACGTTAT 6660

TTATGAGATG GGTTTTTATG ATTAGAGTCC CGCAATTATA CATTTAATAC GCGATAGAAA 6720

ACAAAATATA GCGCGCAAAC TAGGATAAAT TATCGCGCGC GGTGTCATCT ATGTTACTAG 6780

ATCGATCGGG AATTGAGATC TCATATGTCG AGCTCGGGGA TCTCCTTTGC CCCAGAGATC 6840

ACAATGGACG ACTTCCTCTA TCTCTACGAT CTAGTCAGGA AGTTCGACGG AGAAGGTGAC 6900

GATACCATGT TCACCACTGA TAATGAGAAG ATTAGCCTTT TCAATTTCAG AAAGAATGCT 6960

AACCCACAGA TGGTTAGAGA GGCTTACGCA GCAGGTCTCA TCAAGACGAT CTACCCGAGC 7020

AATAATCTCC AGGAGATCAA ATACCTTCCC AAGAAGGTTA AAGATGCAGT CAAAAGATTC 7080

AGGACTAACT GCATCAAGAA CACAGAGAAA GATATATTTC TCAAGATCAG AAGTACTATT 7140

CCAGTATGGA CGATTCAAGG CTTGCTTCAC AAACCAAGGC AAGTAATAGA GATTGGAGTC 7200

TCTAAAAAGG TAGTTCCCAC TGAATCAAAG GCCATGGAGT CAAAGATTCA AATAGAGGAC 7260

CTAACAGAAC TCGCCGTAAA GACTGGCGAA CAGTTCCATC GATGATTGAG ACTTTTCAAC 7320

AAAGGGTAAT ATCCGGAAAC CTCCTCGGAT TCCATTGCCC AGCTATCTGT CACTTTATTG 7380

TGAAGATAGT GGAAAAGGAA GGTGGCTCCT ACAAATGCCA TCATTGCGAT AAAGGAAAGG 7440

CCATCGTTGA AGATGCCTCT GCCGACAGTG GTCCCAAAGA TGGACCCCCA CCCACGAGGA 7500

GCATCGTGGA AAAAGAAGAC GTTCCAACCA CGTCTTCAAA GCAAGTGGAT TGATGTGATA 7560

TCTCCACTGA CGTAAGGGAT GACGCACAAT CCCACTATCC TTCGCAAGAC CCTTCCTCTA 7620

TATAAGGAAG TTCATTTCAT TTGGAGAGGA CACGCTGACA AGCTCGGATC CTTTAGCATG 7680

ATTGAACAAG ATGGATTGCA CGCAGGTTCT CCGGCCGCTT GGGTGGAGAG GCTATTCGGC 7740

TATGACTGGG CACAACAGAC AATCGGCTGC TCTGATGCCG CCGTGTTCCG GCTGTCAGCG 7800

CAGGGGCGCC CGGTTCTTTT TGTCAAGACC GACCTGTCCG GTGCCCTGAA TGAACTGCAG 7860

GACGAGGCAG CGCGGCTATC GTGGCTGGCC ACGACGGGCG TTCCTTGCGC AGCTGTGCTC 7920

GACGTTGTCA CTGAAGCGGG AAGGGACTGG CTGCTATTGG GCGAAGTGCC GGGGCAGGAT 7980

CTCCTGTCAT CTCACCTTGC TCCTGCCGAG AAAGTATCCA TCATGGCTGA TGCAATGCGG 8040

CGGCTGCATA CGCTTGATCC GGCTACCTGC CCATTCGACC ACCAAGCGAA ACATCGCATC 8100

GAGCGAGCAC GTACTCGGAT GGAAGCCGGT CTTGTCGATC AGGATGATCT GGACGAAGAG 8160

CATCAGGGGC TCGCGCCAGC CGAACTGTTC GCCAGGCTCA AGGCGCGCAT GCCCGACGGC 8220

GAGGATCTCG TCGTGACCCA TGGCGATGCC TGCTTGCCGA ATATCATGGT GGAAAATGGC 8280

CGCTTTTCTG GATTCATCGA CTGTGGCCGG CTGGGTGTGG CGGACCGCTA TCAGGACATA 8340

GCGTTGGCTA CCCGTGATAT TGCTGAAGAG CTTGGCGGCG AATGGGCTGA CCGCTTCCTC 8400

GTGCTTTACG GTATCGCCGC TCCCGATTCG CAGCGCATCG CCTTCTATCG CCTTCTTGAC 8460

GAGTTCTTCT GAGCGGGACT CTGGGGTTCG AAATGACCGA CCAAGCGACG CCCAACCTGC 8520

CATCACGAGA TTTCGATTCC ACCGCCGCCT TCTATGAAAG GTTGGGCTTC GGAATCGTTT 8580

TCCGGGACGC CGGCTGGATG ATCCTCCAGC GCGGGGATCT CATGCTGGAG TTCTTCGCCC 8640

ACCCCAACAG AGGTGGATGG ACAGACCCGT TCTTACACCG GACTGGGCGC GGGATAGGAT 8700

ATTCAGATTG GGATGGGATT GAGCTTAAAG CCGGCGCTGA GACCATGCTC AAGGTAGGCA 8760

ATGTCCTCAG CGTCGAGCCC GGCATCTATG TCGAGGGCAT TGGTGGAGCG CGCTTCGGGG 8820

ATACCGTGCT TGTAACTGAG ACCGGATATG AGGCCCTCAC TCCGCTTGAT CTTGGCAAAG 8880

ATATTTGACG CATTTATTAG TATGTGTTAA TTTTCATTTG CAGTGCAGTA TTTTCTATTC 8940

GATCTTTATG TAATTCGTTA CAATTAATAA ATATTCAAAT CAGATTATTG ACTGTCATTT 9000

GTATCAAATC GTGTTTAATG GATATTTTTA TTATAATATT GATGATATCT CAATCAAAAC 9060

GTAGATAATA ATAATATTTA TTTAATATTT TTGCGTCGCA CAGTGAAAAT CTATATGAGA 9120

TTACAAAATA CCGACAACAT TATTTAAGAT ACATAGACAT TAACCCTGAG ACTGTTGGAC 9180

ATCAACGGGT AGATTCCTTC ATGCATAGCA CCTCATTCTT GGGGACAAAA GCACGGTTTG 9240

GCCGTTCCAT TGCTGCACGA ACGAGCTTTG CTATATCCTC GGGTTGGATC ATCTCATCAG 9300

GTCCAATCAA ATTTGTCCAA GAACTCATGT TAGTCGCAAC GAAACCGGGG CATATGGTGC 9360

ACTCTCAGTA CAATCTGCTC TGATGCCGCA TAGTTAAGCC AGCCCCGACA CCCGCCAACA 9420

CCCGCTGACG CGCCCTGACG GGCTTGTCTG CTCCCGGCAT CCGCTTACAG ACAAGCTGTG 9480

ACCGTCTCCG GGAGCTGCAT GTGTCAGAGG TTTTCACCGT CATCACCGAA ACGCGCGAGA 9540

CGAAAGGGCC TCGTGATACG CCTATTTTTA TAGGTTAATG TCATGATAAT AATGGTTTCT 9600

TAGACGTCAG GTGGCACTTT TCGGGGAAAT GTGCGCGGAA CCCCTATTTG TTTATTTTTC 9660

TAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC CCTGATAAAT GCTTCAATAA 9720

TATTGAAAAA GGAAGAGTAT GAGTATTCAA CATTTCCGTG TCGCCCTTAT TCCCTTTTTT 9780

GCGGCATTTT GCCTTCCTGT TTTTGCTCAC CCAGAAACGC TGGTGAAAGT AAAAGATGCT 9840

GAAGATCAGT TGGGTGCACG AGTGGGTTAC ATCGAACTGG ATCTCAACAG CGGTAAGATC 9900

CTTGAGAGTT TTCGCCCCGA AGAACGTTTT CCAATGATGA GCACTTTTAA AGTTCTGCTA 9960

TGTGGCGCGG TATTATCCCG TATTGACGCC GGGCAAGAGC AACTCGGTCG CCGCATACAC 10020

TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC 10080

ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACCATGA GTGATAACAC TGCGGCCAAC 10140

TTACTTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTTGCA CAACATGGGG 10200

GATCATGTAA CTCGCCTTGA TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT ACCAAACGAC 10260

GAGCGTGACA CCACGATGCC TGTAGCAATG GCAACAACGT TGCGCAAACT ATTAACTGGC 10320

GAACTACTTA CTCTAGCTTC CCGGCAACAA TTAATAGACT GGATGGAGGC GGATAAAGTT 10380

GCAGGACCAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT TTATTGCTGA TAAATCTGGA 10440

GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG GGCCAGATGG TAAGCCCTCC 10500

CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG AAATAGACAG 10560

ATCGCTGAGA TAGGTGCCTC ACTGATTAAG CATTGGTAAC TGTCAGACCA AGTTTACTCA 10620

TATATACTTT AGATTGATTT AAAACTTCAT TTTTAATTTA AAAGGATCTA GGTGAAGATC 10680

CTTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTCGTTCCA CTGAGCGTCA 10740

GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT TTTTTCTGCG CGTAATCTGC 10800

TGCTTGCAAA CAAAAAAACC ACCGCTACCA GCGGTGGTTT GTTTGCCGGA TCAAGAGCTA 10860

CCAACTCTTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATACCAAA TACTGTCCTT 10920

CTAGTGTAGC CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC 10980

GCTCTGCTAA TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG TCTTACCGGG 11040

TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT CGGGCTGAAC GGGGGGTTCG 11100

TGCACACAGC CCAGCTTGGA GCGAACGACC TACACCGAAC TGAGATACCT ACAGCGTGAG 11160

CATTGAGAAA GCGCCACGCT TCCCGAAGGG AGAAAGGCGG ACAGGTATCC GGTAAGCGGC 11220

AGGGTCGGAA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCTG GTATCTTTAT 11280

AGTCCTGTCG GGTTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG CTCGTCAGGG 11340

GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GCGGCCTTTT TACGGTTCCT GGCCTTTTGC 11400

TGGCCTTTTG CTCACATGTT CTTTCCTGCG TTATCCCCTG ATTCTGTGGA TAACCGTATT 11460

ACCGCCTTTG AGTGAGCTGA TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTCA 11520

GTGAGCGAGG AAGCGGAAGA GCGCCCAATA CGCAAACCGC CTCTCCCCGC GCGTTGGCCG 11580

ATTCATTAAT GCAGCTGGCA CGACAGGTTT CCCGACTGGA AAGCGGGCAG TGAGCGCAAC 11640

GCAATTAATG TGAGTTAGCT CACTCATTAG GCACCCCAGG CTTTACACTT TATGCTTCCG 11700

GCTCGTATGT TGTGTGGAAT TGTGAGCGGA TAACAATTTC ACACAGGAAA CAGCTATGAC 11760

CATGATTACG CCAAGCTTCC GCGG 11784

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11991 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGGCCCACCA CTGTTGTAAC TTGTAAGCCA CTAGCTCACG TTCTCCATGA GCTCTTCTCT 60

CTGCTGTTTC TTCCTCTGCT AACTGCGTTA TGATATGACG TCGTATAAAT AATCTCACAA 120

TACTTCCTTA TTTTCAGCAT GGCCTCTTTT ATGTTTATTT AACAGTAGCA ACCAACGCCG 180

CTCGATGTTT CCTTCAAGAA ACGGCCACTC ACTATGTGGT GTGCAGAAGA ACAAATGTAA 240

GCAGCTCCTA CAGGTACCAG TAGTCATGTC AGTGTGGAAG CTTTCCAACC AACGCCTCCT 300

TCGAGGAACC TGGTCGTGCT GACATGAATG TAGGCCATGC AAGCACAAGC ACCTAACGCG 360

AATCATCACG ACGCGCCGTG TACTGGGCGT TGGTACATCA CACCCCGCGT TTGACCTGAT 420

CGGAAGCATG CGTGTGTGTT GGCTGCAGGA CCGGCTATAG GTTTCCTGCA TTGGACAGCA 480

GAAGCCAGTC ATGTTAGGCA CTCACGCGCT CCTGCCGTTT GATGAATCAT CCGGTCTTTC 540

GTATTGATCA CTAGTTCACT ACGCTGATAT AGCAAATTTT AAGATGTGAA ACCACGAGAC 600

GAGCGATAAA TCTTAGACGT TACCTATCCA TATGAAGCTT GTGCGAAAAA AAGGCGTGCC 660

GCTGTAGCAT CATTCGTATA CACTTTTGTC CCCAAAGACA GGGATACGAA TCCATGCTCG 720

ACAGAACCCT CCCTTCCCTG CAGATAACGA CACTTAAGTA TAACAAAAGT AGTTGGATTA 780

TTTCAGAAGC AAAATCTCAC TTTTCGCTGG CCTTTTTGTA CTTTGGTTAC TTGAGTTCAG 840

ACAGTGTATG CTATATTGTC ATGTGCTGCG TAAGGTTTAA ATATGGTTCG ACAAATATAT 900

CAGTATATCA CTACTTTGTT ATGGGTGGGG CCTAGCACAA ACTTGATACA GCTAGGATAA 960

AGTTAGAACG ATGACTGATC TACTGTAAAG CGACACCTGT CCTGTTATGG TAGTTTAAGT 1020

CCATTCCTGG ACGACTCCAG ATCCAGGATA TGATGCTGTT ACATAATGCG ATTGTTCACA 1080

ATAAAATTGC ATGATGTTCT TCTACTCTTT AGGCAGTTTT GTTCAACAGG CAAGTTGCAT 1140

AATGCATGTG CATATATGAG CAGCATAATC ATCAATTAAT CATAGGTTCG TCATTTTAGT 1200

TTCACTCCTT CACATTATTC CAGCCCTTGA AGAAAAATGT AGCAGTGCTT GCTGTTTAAT 1260

AAGTGGCAGA GCTGTTTTCA CTCCACCTAC GCTTGTCTAG GACCAAAATT TTAATCTGTC 1320

ACTTTGAGCT AAAACTGAAG CACCAAACCG CTACAAAAGA ACGTAGGAGC TGAATTGTAA 1380

CTTGATGGGA TTACTATAGC AGTTGCTACA GTTCTAGCTA GCTACCTTAT TCTATACGCA 1440

TCACCCTAAC AACCCGGCTG ACTGCTGCAT CTGACCCCAC CGTCCCCTGC TCCAAACCAA 1500

CTCTCCTTTC CTTGCATGCA CTACACCCAC TTCCTGCAGC TATATATACC ACCATATGCC 1560

CATCTTATGA AACCATCCAC AAGAGGAGAA GAAACAATCA ACCAGCAACA CTCTTCTCTT 1620

ATAACATAGT ACAGCGAAGG TAACTCACAG TGCAAAGGTC CGCCTTGTTT CTCCTCTGTC 1680

TCTTGATCTG ACTAATCTTG GTTTATGATT CGTTGAGTAA TTTTGGGGAA AGCTCCTTTG 1740

CTGCTCCACA CATGTCCATT CGAATTTTAC CGTGTTTAGC AAGGGCGAAA AGTTTGCATC 1800

TTGATGATTT AGCTTGACTA TGCGATTGCT TTCCTGGACC CGTGCAGCTG CGCTCGTCGA 1860

CCATGGTCCG TCCTGTAGAA ACCCCAACCC GTGAAATCAA AAAACTCGAC GGCCTGTGGG 1920

CATTCAGTCT GGATCGCGAA AACTGTGGAA TTGATCAGCG TTGGTGGGAA AGCGCGTTAC 1980

AAGAAAGCCG GGCAATTGCT GTGCCAGGCA GTTTTAACGA TCAGTTCGCC GATGCAGATA 2040

TTCGTAATTA TGCGGGCAAC GTCTGGTATC AGCGCGAAGT CTTTATACCG AAAGGTTGGG 2100

CAGGCCAGCG TATCGTGCTG CGTTTCGATG CGGTCACTCA TTACGGCAAA GTGTGGGTCA 2160

ATAATCAGGA AGTGATGGAG CATCAGGGCG GCTATACGCC ATTTGAAGCC GATGTCACGC 2220

CGTATGTTAT TGCCGGGAAA AGTGTACGTA TCACCGTTTG TGTGAACAAC GAACTGAACT 2280

GGCAGACTAT CCCGCCGGGA ATGGTGATTA CCGACGAAAA CGGCAAGAAA AAGCAGTCTT 2340

ACTTCCATGA TTTCTTTAAC TATGCCGGAA TCCATCGCAG CGTAATGCTC TACACCACGC 2400

CGAACACCTG GGTGGACGAT ATCACCGTGG TGACGCATGT CGCGCAAGAC TGTAACCACG 2460

CGTCTGTTGA CTGGCAGGTG GTGGCCAATG GTGATGTCAG CGTTGAACTG CGTGATGCGG 2520

ATCAACAGGT GGTTGCAACT GGACAAGGCA CTAGCGGGAC TTTGCAAGTG GTGAATCCGC 2580

ACCTCTGGCA ACCGGGTGAA GGTTATCTCT ATGAACTGTG CGTCACAGCC AAAAGCCAGA 2640

CAGAGTGTGA TATCTACCCG CTTCGCGTCG GCATCCGGTC AGTGGCAGTG AAGGGCGAAC 2700

AGTTCCTGAT TAACCACAAA CCGTTCTACT TTACTGGCTT TGGTCGTCAT GAAGATGCGG 2760

ACTTACGTGG CAAAGGATTC GATAACGTGC TGATGGTGCA CGACCACGCA TTAATGGACT 2820

GGATTGGGGC CAACTCCTAC CGTACCTCGC ATTACCCTTA CGCTGAAGAG ATGCTCGACT 2880

GGGCAGATGA ACATGGCATC GTGGTGATTG ATGAAACTGC TGCTGTCGGC TTTAACCTCT 2940

CTTTAGGCAT TGGTTTCGAA GCGGGCAACA AGCCGAAAGA ACTGTACAGC GAAGAGGCAG 3000

TCAACGGGGA AACTCAGCAA GCGCACTTAC AGGCGATTAA AGAGCTGATA GCGCGTGACA 3060

AAAACCACCC AAGCGTGGTG ATGTGGAGTA TTGCCAACGA ACCGGATACC CGTCCGCAAG 3120

TGCACGGGAA TATTTCGCCA CTGGCGGAAG CAACGCGTAA ACTCGACCCG ACGCGTCCGA 3180

TCACCTGCGT CAATGTAATG TTCTGCGACG CTCACACCGA TACCATCAGC GATCTCTTTG 3240

ATGTGCTGTG CCTGAACCGT TATTACGGAT GGTATGTCCA AAGCGGCGAT TTGGAAACGG 3300

CAGAGAAGGT ACTGGAAAAA GAACTTCTGG CCTGGCAGGA GAAACTGCAT CAGCCGATTA 3360

TCATCACCGA ATACGGCGTG GATACGTTAG CCGGGCTGCA CTCAATGTAC ACCGACATGT 3420

GGAGTGAAGA GTATCAGTGT GCATGGCTGG ATATGTATCA CCGCGTCTTT GATCGCGTCA 3480

GCGCCGTCGT CGGTGAACAG GTATGGAATT TCGCCGATTT TGCGACCTCG CAAGGCATAT 3540

TGCGCGTTGG CGGTAACAAG AAAGGGATCT TCACTCGCGA CCGCAAACCG AAGTCGGCGG 3600

CTTTTCTGCT GCAAAAACGC TGGACTGGCA TGAACTTCGG TGAAAAACCG CAGCAGGGAG 3660

GCAAACAATG AATCAACAAC TCTCCTGGCG CACCATCGTC GGCTACAGCC TCGGTGGGGA 3720

ATTGGAGCTC GAATTTCCCC GATCGTTCAA ACATTTGGCA ATAAAGTTTC TTAAGATTGA 3780

ATCCTGTTGC CGGTCTTGCG ATGATTATCA TATAATTTCT GTTGAATTAC GTTAAGCATG 3840

TAATAATTAA CATGTAATGC ATGACGTTAT TTATGAGATG GGTTTTTATG ATTAGAGTCC 3900

CGCAATTATA CATTTAATAC GCGATAGAAA ACAAAATATA GCGCGCAAAC TAGGATAAAT 3960

TATCGCGCGC GGTGTCATCT ATGTTACTAG ATCGATCGGG AATTAAGCTT AGATCTGCAT 4020

GGGTGGAGAC TTTTCAACAA AGGGTAATAT CCGGAAACCT CCTCGGATTC CATTGCCCAG 4080

CTATCTGTCA CTTTATTGTG AAGATAGTGG AAAAGGAAGG TGGCTCCTAC AAATGCCATC 4140

ATTGCGATAA AGGAAAGGCC ATCGTTGAAG ATGCCTCTGC CGACAGTGGT CCCAAAGATG 4200

GACCCCCACC CACGAGGAGC ATCGTGGAAA AAGAAGACGT TCCAACCACG TCTTCAAAGC 4260

AAGTGGATTG ATGTGATCAT CGATGGAGAC TTTTCAACAA AGGGTAATAT CCGGAAACCT 4320

CCTCGGATTC CATTGCCCAG CTATCTGTCA CTTTATTGTG AAGATAGTGG AAAAGGAAGG 4380

TGGCTCCTAC AAATGCCATC ATTGCGATAA AGGAAAGGCC ATCGTTGAAG ATGCCTCTGC 4440

CGACAGTGGT CCCAAAGATG GACCCCCACC CACGAGGAGC ATCGTGGAAA AAGAAGACGT 4500

TCCAACCACG TCTTCAAAGC AAGTGGATTG ATGTGATATC TCCACTGACG TAAGGGATGA 4560

CGCACAATCC CACTATCCTT CGCAAGACCC TTCCTCTATA TAAGGAAGTT CATTTCATTT 4620

GGAGAGAACA CGGGGGACTC TAGAGGATCC AGCTGAAGGC TCGACAAGGC AGTCCACGGA 4680

GGAGCTGATA TTTGGTGGAC AAGCTGTGGA TAGGAGCAAC CCTATCCCTA ATATACCAGC 4740

ACCACCAAGT CAGGGCAATC CCCAGATCAA GTGCAAAGGT CCGCCTTGTT TCTCCTCTGT 4800

CTCTTGATCT GACTAATCTT GGTTTATGAT TCGTTGAGTA ATTTTGGGGA AAGCTCCTTT 4860

GCTGCTCCAC ACATGTCCAT TCGAATTTTA CCGTGTTTAG CAAGGGCGAA AAGTTTGCAT 4920

CTTGATGATT TAGCTTGACT ATGCGATTGC TTTCCTGGAC CCGTGCAGCT GCGCTCGGAT 4980

CTGGGGCCAT TTGTTCCAGG CACGGGATAA GCATTCAGCC ATGGCAGACG CCAAAAACAT 5040

AAAGAAAGGC CCGGCGCCAT TCTATCCTCT AGAGGATGGA ACCGCTGGAG AGCAACTGCA 5100

TAAGGCTATG AAGAGATACG CCCTGGTTCC TGGAACAATT GCTTTTACAG ATGCACATAT 5160

CGAGGTGAAC ATCACGTACG CGGAATACTT CGAAATGTCC GTTCGGTTGG CAGAAGCTAT 5220

GAAACGATAT GGGCTGAATA CAAATCACAG AATCGTCGTA TGCAGTGAAA ACTCTCTTCA 5280

ATTCTTTATG CCGGTGTTGG GCGCGTTATT TATCGGAGTT GCAGTTGCGC CCGCGAACGA 5340

CATTTATAAT GAACGTGAAT TGCTCAACAG TATGAACATT TCGCAGCCTA CCGTAGTGTT 5400

TGTTTCCAAA AAGGGGTTGC AAAAAATTTT GAACGTGCAA AAAAAATTAC CAATAATCCA 5460

GAAAATTATT ATCATGGATT CTAAAACGGA TTACCAGGGA TTTCAGTCGA TGTACACGTT 5520

CGTCACATCT CATCTACCTC CCGGTTTTAA TGAATACGAT TTTGTACCAG AGTCCTTTGA 5580

TCGTGACAAA ACAATTGCAC TGATAATGAA TTCCTCTGGA TCTACTGGGT TACCTAAGGG 5640

TGTGGCCCTT CCGCATAGAA CTGCCTGCGT CAGATTCTCG CATGCCAGAG ATCCTATTTT 5700

TGGCAATCAA ATCATTCCGG ATACTGCGAT TTTAAGTGTT GTTCCATTCC ATCACGGTTT 5760

TGGAATGTTT ACTACACTCG GATATTTGAT ATGTGGATTT CGAGTCGTCT TAATGTATAG 5820

ATTTGAAGAA GAGCTGTTTT TACGATCCCT TCAGGATTAC AAAATTCAAA GTGCGTTGCT 5880

AGTACCAACC CTATTTTCAT TCTTCGCCAA AAGCACTCTG ATTGACAAAT ACGATTTATC 5940

TAATTTACAC GAAATTGCTT CTGGGGGCGC ACCTCTTTCG AAAGAAGTCG GGGAAGCGGT 6000

TGCAAAACGC TTCCATCTTC CAGGGATACG ACAAGGATAT GGGCTCACTG AGACTACATC 6060

AGCTATTCTG ATTACACCCG AGGGGGATGA TAAACCGGGC GCGGTCGGTA AAGTTGTTCC 6120

ATTTTTTGAA GCGAAGGTTG TGGATCTGGA TACCGGGAAA ACGCTGGGCG TTAATCAGAG 6180

AGGCGAATTA TGTGTCAGAG GACCTATGAT TATGTCCGGT TATGTAAACA ATCCGGAAGC 6240

GACCAACGCC TTGATTGACA AGGATGGATG GCTACATTCT GGAGACATAG CTTACTGGGA 6300

CGAAGACGAA CACTTCTTCA TAGTTGACCG CTTGAAGTCT TTAATTAAAT ACAAAGGATA 6360

TCAGGTGGCC CCCGCTGAAT TGGAATCGAT ATTGTTACAA CACCCCAACA TCTTCGACGC 6420

GGGCGTGGCA GGTCTTCCCG ACGATGACGC CGGTGAACTT CCCGCCGCCG TTGTTGTTTT 6480

GGAGCACGGA AAGACGATGA CGGAAAAAGA GATCGTGGAT TACGTCGCCA GTCAAGTAAC 6540

AACCGCGAAA AAGTTGCGCG GAGGAGTTGT GTTTGTGGAC GAAGTACCGA AAGGTCTTAC 6600

CGGAAAACTC GACGCAAGAA AAATCAGAGA GATCCTCATA AAGGCCAAGA AGGGCGGAAA 6660

GTCCAAATTG TAAAATGTAA CTGTATTCAG CGATGACGAA ATTCTTAGCT ATTGTAATCA 6720

GATCCGCGAA TTTCCCCGAT CGTTCAAACA TTTGGCAATA AAGTTTCTTA AGATTGAATC 6780

CTGTTGCCGG TCTTGCGATG ATTATCATAT AATTTCTGTT GAATTACGTT AAGCATGTAA 6840

TAATTAACAT GTAATGCATG ACGTTATTTA TGAGATGGGT TTTTATGATT AGAGTCCCGC 6900

AATTATACAT TTAATACGCG ATAGAAAACA AAATATAGCG CGCAAACTAG GATAAATTAT 6960

CGCGCGCGGT GTCATCTATG TTACTAGATC GATCGGGAAT TGAGATCTCA TATGTCGAGC 7020

TCGGGGATCT CCTTTGCCCC AGAGATCACA ATGGACGACT TCCTCTATCT CTACGATCTA 7080

GTCAGGAAGT TCGACGGAGA AGGTGACGAT ACCATGTTCA CCACTGATAA TGAGAAGATT 7140

AGCCTTTTCA ATTTCAGAAA GAATGCTAAC CCACAGATGG TTAGAGAGGC TTACGCAGCA 7200

GGTCTCATCA AGACGATCTA CCCGAGCAAT AATCTCCAGG AGATCAAATA CCTTCCCAAG 7260

AAGGTTAAAG ATGCAGTCAA AAGATTCAGG ACTAACTGCA TCAAGAACAC AGAGAAAGAT 7320

ATATTTCTCA AGATCAGAAG TACTATTCCA GTATGGACGA TTCAAGGCTT GCTTCACAAA 7380

CCAAGGCAAG TAATAGAGAT TGGAGTCTCT AAAAAGGTAG TTCCCACTGA ATCAAAGGCC 7440

ATGGAGTCAA AGATTCAAAT AGAGGACCTA ACAGAACTCG CCGTAAAGAC TGGCGAACAG 7500

TTCCATCGAT GATTGAGACT TTTCAACAAA GGGTAATATC CGGAAACCTC CTCGGATTCC 7560

ATTGCCCAGC TATCTGTCAC TTTATTGTGA AGATAGTGGA AAAGGAAGGT GGCTCCTACA 7620

AATGCCATCA TTGCGATAAA GGAAAGGCCA TCGTTGAAGA TGCCTCTGCC GACAGTGGTC 7680

CCAAAGATGG ACCCCCACCC ACGAGGAGCA TCGTGGAAAA AGAAGACGTT CCAACCACGT 7740

CTTCAAAGCA AGTGGATTGA TGTGATATCT CCACTGACGT AAGGGATGAC GCACAATCCC 7800

ACTATCCTTC GCAAGACCCT TCCTCTATAT AAGGAAGTTC ATTTCATTTG GAGAGGACAC 7860

GCTGACAAGC TCGGATCCTT TAGCATGATT GAACAAGATG GATTGCACGC AGGTTCTCCG 7920

GCCGCTTGGG TGGAGAGGCT ATTCGGCTAT GACTGGGCAC AACAGACAAT CGGCTGCTCT 7980

GATGCCGCCG TGTTCCGGCT GTCAGCGCAG GGGCGCCCGG TTCTTTTTGT CAAGACCGAC 8040

CTGTCCGGTG CCCTGAATGA ACTGCAGGAC GAGGCAGCGC GGCTATCGTG GCTGGCCACG 8100

ACGGGCGTTC CTTGCGCAGC TGTGCTCGAC GTTGTCACTG AAGCGGGAAG GGACTGGCTG 8160

CTATTGGGCG AAGTGCCGGG GCAGGATCTC CTGTCATCTC ACCTTGCTCC TGCCGAGAAA 8220

GTATCCATCA TGGCTGATGC AATGCGGCGG CTGCATACGC TTGATCCGGC TACCTGCCCA 8280

TTCGACCACC AAGCGAAACA TCGCATCGAG CGAGCACGTA CTCGGATGGA AGCCGGTCTT 8340

GTCGATCAGG ATGATCTGGA CGAAGAGCAT CAGGGGCTCG CGCCAGCCGA ACTGTTCGCC 8400

AGGCTCAAGG CGCGCATGCC CGACGGCGAG GATCTCGTCG TGACCCATGG CGATGCCTGC 8460

TTGCCGAATA TCATGGTGGA AAATGGCCGC TTTTCTGGAT TCATCGACTG TGGCCGGCTG 8520

GGTGTGGCGG ACCGCTATCA GGACATAGCG TTGGCTACCC GTGATATTGC TGAAGAGCTT 8580

GGCGGCGAAT GGGCTGACCG CTTCCTCGTG CTTTACGGTA TCGCCGCTCC CGATTCGCAG 8640

CGCATCGCCT TCTATCGCCT TCTTGACGAG TTCTTCTGAG CGGGACTCTG GGGTTCGAAA 8700

TGACCGACCA AGCGACGCCC AACCTGCCAT CACGAGATTT CGATTCCACC GCCGCCTTCT 8760

ATGAAAGGTT GGGCTTCGGA ATCGTTTTCC GGGACGCCGG CTGGATGATC CTCCAGCGCG 8820

GGGATCTCAT GCTGGAGTTC TTCGCCCACC CCAACAGAGG TGGATGGACA GACCCGTTCT 8880

TACACCGGAC TGGGCGCGGG ATAGGATATT CAGATTGGGA TGGGATTGAG CTTAAAGCCG 8940

GCGCTGAGAC CATGCTCAAG GTAGGCAATG TCCTCAGCGT CGAGCCCGGC ATCTATGTCG 9000

AGGGCATTGG TGGAGCGCGC TTCGGGGATA CCGTGCTTGT AACTGAGACC GGATATGAGG 9060

CCCTCACTCC GCTTGATCTT GGCAAAGATA TTTGACGCAT TTATTAGTAT GTGTTAATTT 9120

TCATTTGCAG TGCAGTATTT TCTATTCGAT CTTTATGTAA TTCGTTACAA TTAATAAATA 9180

TTCAAATCAG ATTATTGACT GTCATTTGTA TCAAATCGTG TTTAATGGAT ATTTTTATTA 9240

TAATATTGAT GATATCTCAA TCAAAACGTA GATAATAATA ATATTTATTT AATATTTTTG 9300

CGTCGCACAG TGAAAATCTA TATGAGATTA CAAAATACCG ACAACATTAT TTAAGATACA 9360

TAGACATTAA CCCTGAGACT GTTGGACATC AACGGGTAGA TTCCTTCATG CATAGCACCT 9420

CATTCTTGGG GACAAAAGCA CGGTTTGGCC GTTCCATTGC TGCACGAACG AGCTTTGCTA 9480

TATCCTCGGG TTGGATCATC TCATCAGGTC CAATCAAATT TGTCCAAGAA CTCATGTTAG 9540

TCGCAACGAA ACCGGGGCAT ATGGTGCACT CTCAGTACAA TCTGCTCTGA TGCCGCATAG 9600

TTAAGCCAGC CCCGACACCC GCCAACACCC GCTGACGCGC CCTGACGGGC TTGTCTGCTC 9660

CCGGCATCCG CTTACAGACA AGCTGTGACC GTCTCCGGGA GCTGCATGTG TCAGAGGTTT 9720

TCACCGTCAT CACCGAAACG CGCGAGACGA AAGGGCCTCG TGATACGCCT ATTTTTATAG 9780

GTTAATGTCA TGATAATAAT GGTTTCTTAG ACGTCAGGTG GCACTTTTCG GGGAAATGTG 9840

CGCGGAACCC CTATTTGTTT ATTTTTCTAA ATACATTCAA ATATGTATCC GCTCATGAGA 9900

CAATAACCCT GATAAATGCT TCAATAATAT TGAAAAAGGA AGAGTATGAG TATTCAACAT 9960

TTCCGTGTCG CCCTTATTCC CTTTTTTGCG GCATTTTGCC TTCCTGTTTT TGCTCACCCA 10020

GAAACGCTGG TGAAAGTAAA AGATGCTGAA GATCAGTTGG GTGCACGAGT GGGTTACATC 10080

GAACTGGATC TCAACAGCGG TAAGATCCTT GAGAGTTTTC GCCCCGAAGA ACGTTTTCCA 10140

ATGATGAGCA CTTTTAAAGT TCTGCTATGT GGCGCGGTAT TATCCCGTAT TGACGCCGGG 10200

CAAGAGCAAC TCGGTCGCCG CATACACTAT TCTCAGAATG ACTTGGTTGA GTACTCACCA 10260

GTCACAGAAA AGCATCTTAC GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA 10320

ACCATGAGTG ATAACACTGC GGCCAACTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG 10380

CTAACCGCTT TTTTGCACAA CATGGGGGAT CATGTAACTC GCCTTGATCG TTGGGAACCG 10440

GAGCTGAATG AAGCCATACC AAACGACGAG CGTGACACCA CGATGCCTGT AGCAATGGCA 10500

ACAACGTTGC GCAAACTATT AACTGGCGAA CTACTTACTC TAGCTTCCCG GCAACAATTA 10560

ATAGACTGGA TGGAGGCGGA TAAAGTTGCA GGACCACTTC TGCGCTCGGC CCTTCCGGCT 10620

GGCTGGTTTA TTGCTGATAA ATCTGGAGCC GGTGAGCGTG GGTCTCGCGG TATCATTGCA 10680

GCACTGGGGC CAGATGGTAA GCCCTCCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG 10740

GCAACTATGG ATGAACGAAA TAGACAGATC GCTGAGATAG GTGCCTCACT GATTAAGCAT 10800

TGGTAACTGT CAGACCAAGT TTACTCATAT ATACTTTAGA TTGATTTAAA ACTTCATTTT 10860

TAATTTAAAA GGATCTAGGT GAAGATCCTT TTTGATAATC TCATGACCAA AATCCCTTAA 10920

CGTGAGTTTT CGTTCCACTG AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA 10980

GATCCTTTTT TTCTGCGCGT AATCTGCTGC TTGCAAACAA AAAAACCACC GCTACCAGCG 11040

GTGGTTTGTT TGCCGGATCA AGAGCTACCA ACTCTTTTTC CGAAGGTAAC TGGCTTCAGC 11100

AGAGCGCAGA TACCAAATAC TGTCCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG 11160

AACTCTGTAG CACCGCCTAC ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC 11220

AGTGGCGATA AGTCGTGTCT TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG 11280

CAGCGGTCGG GCTGAACGGG GGGTTCGTGC ACACAGCCCA GCTTGGAGCG AACGACCTAC 11340

ACCGAACTGA GATACCTACA GCGTGAGCAT TGAGAAAGCG CCACGCTTCC CGAAGGGAGA 11400

AAGGCGGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG GAGAGCGCAC GAGGGAGCTT 11460

CCAGGGGGAA ACGCCTGGTA TCTTTATAGT CCTGTCGGGT TTCGCCACCT CTGACTTGAG 11520

CGTCGATTTT TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG 11580

GCCTTTTTAC GGTTCCTGGC CTTTTGCTGG CCTTTTGCTC ACATGTTCTT TCCTGCGTTA 11640

TCCCCTGATT CTGTGGATAA CCGTATTACC GCCTTTGAGT GAGCTGATAC CGCTCGCCGC 11700

AGCCGAACGA CCGAGCGCAG CGAGTCAGTG AGCGAGGAAG CGGAAGAGCG CCCAATACGC 11760

AAACCGCCTC TCCCCGCGCG TTGGCCGATT CATTAATGCA GCTGGCACGA CAGGTTTCCC 11820

GACTGGAAAG CGGGCAGTGA GCGCAACGCA ATTAATGTGA GTTAGCTCAC TCATTAGGCA 11880

CCCCAGGCTT TACACTTTAT GCTTCCGGCT CGTATGTTGT GTGGAATTGT GAGCGGATAA 11940

CAATTTCACA CAGGAAACAG CTATGACCAT GATTACGCCA AGCTTCCGCG G 11991

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ACGTACGTAC GGGCCCACCA CTGTTGTAAC TTGTAAGCC 39

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AGGCGGACCT TTGCACTGTG AGTTACCTTC GC 32

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CTCTGTCGAC GAGCGCAGCT GCACGGGTC 29

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCGAAGGTAA CTCACAGTGC AAAGGTCCGC CT 32

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9299 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GAGCTCCACC GCGGTGGCGG CCGCTCTAGA ACTAGTGGAT CCGTCGACCA TGGCCAGTTG 60

CCGGTGGAGC AGGTAAAAAC ACCGTAGCGT AGCAGCCAGG CGGAAGCAGA CGCACAGCAC 120

AGGTTGGTTA TGATAGTCAG CCGGGCCACA TGTGTGTAGT TGGTACACTG ATACGCTTAC 180

ACTGTCTCTC CTTTCTTTTT TATTTGTCAC CTTTGGTCGA GCTTACATAA TTGTGTGACT 240

AAAAAAAGGT CACTTCATTC AGAAATTTAG GGTTGTGGGA ATTTTGGATT TTATTGTGTC 300

TGTATAGAGT AGCTATAGCT AGCTAGCTAG ATGTGATGTT AATAATTATG ACGATGAGAT 360

TGGCCCGCTT GGCCGCTTGC ATTGTCTCCC TAGCTCAATA ATGTTTTGAG TTTGTCTTGC 420

CTTTCTTTCA GCTCTAACAA ATTGGAGTAG GGATGACTGA GATACATATA TAAAAGCGAA 480

AACCGCTGCT CTCTGTTAAT TATTGCACAT CACACATAGG CCAAGCCTTA AGGACAATCA 540

ACTAAGGATG GTAATAACTA AGGCTAGTGA GGTCGAACTA GGGATGTTAA TATACTCTAG 600

ATTTTAGACT ATAAAATTTA AGGATCGAAT CAGATTAGTA TCGAACTATA TTTATATTCA 660

TTTCTAAACT AAATTAATTA AGCACCCTAA ATTATTGTGA TGAAGAGACA TTTCGATCGT 720

GATCCATTAT TACTCCTTGG TCAAACTAAT CTCGTTTTAT GTCACTATTT CATCATCTTT 780

TTTGCGAACG GGTTTATAGC CCGTGTTCCA TTATGAGGAC ATGAACGGTT TAAACAAAGT 840

TACATATCAT CCCAGCTAGC TACCTAGATT GGAAGCATGG GTTCGGTATA TATATATAGT 900

TTATATATTT GGTATATATA TATATATATA TATATATATA TATATATCAC ACGTCAGCTT 960

ATATTACGTA AAGTGGGGTT AGTTTTCAAG AAGCGTGGGA CCAGTCACCT CTGCAGTCTG 1020

ACCTTGGCTT CAGCTTCGAC AGCAAACAGT CATCTCTTGG AAGCTAAGGA CAGTCTCCAA 1080

CAGTCAACAA AGCAGCGGTC TGCTTGTAGT TCTCCCTTGC ACGACCAGCT ATATCTAGCA 1140

TCATAACAAC GGTAAGATCA TCTCTAGCAC GACAAACTTA GTTTAATTAA TTATGTCTAA 1200

TCCGTTGTTG TTAGCTTAAA CTTTCTAGCC TCCTATGCTA AGAGAGTTCT CTAGTTCTAC 1260

TCAGGTGGAT TGATATATAA ATTGGGAATC TTCTAGGCGT CACAAGGTAT GGTACACATC 1320

AATCAATGAA CGGACAAAGC AACGGTAAGA TCCGACCCAG TAAAAGTAAT AGCGTTAGGG 1380

CATGTACAAC CTAGACACTG ATGCACAGTA CTCCAAGTAT AAGACACAAC TAAAACACAA 1440

CATAATAATA CAGTGGTTAT ATCTAAAACA TGTGTCTTAC CATATTCATT GTACCAATTA 1500

GAACATTTAA TAAATTAAAG TGACCAATCA GCTAGCCTCC TGTCTCGAAC ATAGAGCTAA 1560

GACATTGTGT CTTCGTCAAG ATACATGTCT TAAGTTTTTT TATATTCACT CCCAAAGACA 1620

CACTCTAAGA CACAACGTAA CACACCCATT GTACATGCTC TTAACCTAAG TTATCATGGA 1680

TGACCACGCG TGGCAATTAA AAAAATAATT TTTGCCTCCT AAAACCTCTT TCTTAATTGG 1740

TTCTTGCTTG CAAATCACCA GCGAACCCAT ATGAAAGGAT GCTCAAAATC TGGCCACCGC 1800

ATCAGGGTTG GTGAATGCAA VGTAAAAAAT AATGCATAAA TCAGCTCTCT GATCAGTTAT 1860

ATAATCGTGC CTTTTAATTA TTCATGCCAG CTTTATCTGA CTCACGAAAT CATTGATAAA 1920

TTATTCCTCA GCTGTATTAG AAAGAGCAGT GTTGTTTAAC TTGGAAAGTG ATGTGGAAGC 1980

GTGTGATTGC GGTTGAGCTT GTATAGGAGT AAAATGAGGA ACAGTAGGAA AATAATTTTT 2040

TCGGATTAAA ACCGGTTGTT TGGACTGCGG CAGATACAAT TCATAGAGAT AAAAACACCG 2100

TAGAAGTATT AGAAGCCGAT AAAGATTAAA CCCAAATGAA CGAACAGGCT AAACAAATCC 2160

GGCGCCTCAA AAGTCAAGAG CAGGTACTGG GCTGTCTTGC ACACGTCGCT TTTTGTCTCC 2220

CCCTGGCCCC TGGGTGAGAG TAGTAGGGAT GCTAAAGTTT GCTTTCTCTT TTTGAGGCAT 2280

GTGATAGGCT CTTGTTAGTT GCTAGGGCTA TGTTTATAAT ATTTGCGCTT TTACCTATGT 2340

ACGTAAGAAC CGGATGGAAT AATGCTATGC AGGAACCAAT TATGTTTGGT CGAAATATAT 2400

AGTGACCTAT CATAATGTTA TCCCTGTTCA TGTACCTAGG TGGCTAATGA TATACGGCAT 2460

ATGAATACAG TAATCATCCA AGCACGTAAA AACTCGCTAG ACGTTTATGC CTGCTAGCCT 2520

GCTGGGTGTG TAGACTGGAG TACTGGACAA ACATCGCAAT ACAGAGGTAC AGTATTTGTC 2580

TAGACAATGA TATACATAGA TAAAAACCAC TGTTGTAACT TGTAAGCCAC TAGCTCACGT 2640

TCTCCATGAG CTCTTCTCTC TGCTGTTTCT TCCTCTGCTA ACTGCGTTAT GATATGACGT 2700

CGTATAAATA ATCTCACAAT ACTTCCTTAT TTTCAGCATG GCCTCTTTTA TGTTTATTTA 2760

ACAGTAGCAA CCAACGCCGC TCGATGTTTC CTTCAAGAAA CGGCCACTCA CTATGTGGTG 2820

TGCAGAAGAA CAAATGTAAG CAGCTCCTAC AGGTACCAGT AGTCATGTCA GTGTGGAAGC 2880

TTTCCAACCA ACGCCTCCTT CGAGGAACCT GGTCGTGCTG ACATGAATGT AGGCCATGCA 2940

AGCACAAGCA CCTAACGCGA ATCATCACGA CGCGCCGTGT ACTGGGCGTT GGTACATCAC 3000

ACCCCGCGTT TGACCTGATC GGAAGCATGC GTGTGTGTTG GCTGCAGGAC CGGCTATAGG 3060

TTTCCTGCAT TGGACAGCAG AAGCCAGTCA TGTTAGGCAC TCACGCGCTC CTGCCGTTTG 3120

ATGAATCATC CGGTCTTTCG TATTGATCAC TAGTTCACTA CGCTGATATA GCAAATTTTA 3180

AGATGTGAAA CCACGAGACG AGCGATAAAT CTTAGACGTT ACCTATCCAT ATGAAGCTTG 3240

TGCGAAAAAA AGGCGTGCCG CTGTAGCATC ATTCGTATAC ACTTTTGTCC CCAAAGACAG 3300

GGATACGAAT CCATGCTCGA CAGAACCCTC CCTTCCCTGC AGATAACGAC ACTTAAGTAT 3360

AACAAAAGTA GTTGGATTAT TTCAGAAGCA AAATCTCACT TTTCGCTGGC CTTTTTGTAC 3420

TTTGGTTACT TGAGTTCAGA CAGTGTATGC TATATTGTCA TGTGCTGCGT AAGGTTTAAA 3480

TATGGTTCGA CAAATATATC AGTATATCAC TACTTTGTTA TGGGTGGGGC CTAGCACAAA 3540

CTTGATACAG CTAGGATAAA GTTAGAACGA TGACTGATCT ACTGTAAAGC GACACCTGTC 3600

CTGTTATGGT AGTTTAAGTC CATTCCTGGA CGACTCCAGA TCCAGGATAT GATGCTGTTA 3660

CATAATGCGA TTGTTCACAA TAAAATTGCA TGATGTTCTT CTACTCTTTA GGCAGTTTTG 3720

TTCAACAGGC AAGTTGCATA ATGCATGTGC ATATATGAGC AGCATAATCA TCAATTAATC 3780

ATAGGTTCGT CATTTTAGTT TCACTCCTTC ACATTATTCC AGCCCTTGAA GAAAAATGTA 3840

GCAGTGCTTG CTGTTTAATA AGTGGCAGAG CTGTTTTCAC TCCACCTACG CTTGTCTAGG 3900

ACCAAAATTT TAATCTGTCA CTTTGAGCTA AAACTGAAGC ACCAAACCGC TACAAAAGAA 3960

CGTAGGAGCT GAATTGTAAC TTGATGGGAT TACTATAGCA GTTGCTACAG TTCTAGCTAG 4020

CTACCTTATT CTATACGCAT CACCCTAACA ACCCGGCTGA CTGCTGCATC TGACCCCACC 4080

GTCCCCTGCT CCAAACCAAC TCTCCTTTCC TTGCATGCAC TACACCCACT TCCTGCAGCT 4140

ATATATACCA CCATATGCCC ATCTTATGAA ACCATCCACA AGAGGAGAAG AAACAATCAA 4200

CCAGCAACAC TCTTCTCTTA TAACATAGTA CAGCGAAGGT AACTCACATG GCAACTTCCA 4260

TGGTCCGTCC TGTAGAAACC CCAACCCGTG AAATCAAAAA ACTCGACGGC CTGTGGGCAT 4320

TCAGTCTGGA TCGCGAAAAC TGTGGAATTG ATCAGCGTTG GTGGGAAAGC GCGTTACAAG 4380

AAAGCCGGGC AATTGCTGTG CCAGGCAGTT TTAACGATCA GTTCGCCGAT GCAGATATTC 4440

GTAATTATGC GGGCAACGTC TGGTATCAGC GCGAAGTCTT TATACCGAAA GGTTGGGCAG 4500

GCCAGCGTAT CGTGCTGCGT TTCGATGCGG TCACTCATTA CGGCAAAGTG TGGGTCAATA 4560

ATCAGGAAGT GATGGAGCAT CAGGGCGGCT ATACGCCATT TGAAGCCGAT GTCACGCCGT 4620

ATGTTATTGC CGGGAAAAGT GTACGTATCA CCGTTTGTGT GAACAACGAA CTGAACTGGC 4680

AGACTATCCC GCCGGGAATG GTGATTACCG ACGAAAACGG CAAGAAAAAG CAGTCTTACT 4740

TCCATGATTT CTTTAACTAT GCCGGAATCC ATCGCAGCGT AATGCTCTAC ACCACGCCGA 4800

ACACCTGGGT GGACGATATC ACCGTGGTGA CGCATGTCGC GCAAGACTGT AACCACGCGT 4860

CTGTTGACTG GCAGGTGGTG GCCAATGGTG ATGTCAGCGT TGAACTGCGT GATGCGGATC 4920

AACAGGTGGT TGCAACTGGA CAAGGCACTA GCGGGACTTT GCAAGTGGTG AATCCGCACC 4980

TCTGGCAACC GGGTGAAGGT TATCTCTATG AACTGTGCGT CACAGCCAAA AGCCAGACAG 5040

AGTGTGATAT CTACCCGCTT CGCGTCGGCA TCCGGTCAGT GGCAGTGAAG GGCGAACAGT 5100

TCCTGATTAA CCACAAACCG TTCTACTTTA CTGGCTTTGG TCGTCATGAA GATGCGGACT 5160

TACGTGGCAA AGGATTCGAT AACGTGCTGA TGGTGCACGA CCACGCATTA ATGGACTGGA 5220

TTGGGGCCAA CTCCTACCGT ACCTCGCATT ACCCTTACGC TGAAGAGATG CTCGACTGGG 5280

CAGATGAACA TGGCATCGTG GTGATTGATG AAACTGCTGC TGTCGGCTTT AACCTCTCTT 5340

TAGGCATTGG TTTCGAAGCG GGCAACAAGC CGAAAGAACT GTACAGCGAA GAGGCAGTCA 5400

ACGGGGAAAC TCAGCAAGCG CACTTACAGG CGATTAAAGA GCTGATAGCG CGTGACAAAA 5460

ACCACCCAAG CGTGGTGATG TGGAGTATTG CCAACGAACC GGATACCCGT CCGCAAGTGC 5520

ACGGGAATAT TTCGCCACTG GCGGAAGCAA CGCGTAAACT CGACCCGACG CGTCCGATCA 5580

CCTGCGTCAA TGTAATGTTC TGCGACGCTC ACACCGATAC CATCAGCGAT CTCTTTGATG 5640

TGCTGTGCCT GAACCGTTAT TACGGATGGT ATGTCCAAAG CGGCGATTTG GAAACGGCAG 5700

AGAAGGTACT GGAAAAAGAA CTTCTGGCCT GGCAGGAGAA ACTGCATCAG CCGATTATCA 5760

TCACCGAATA CGGCGTGGAT ACGTTAGCCG GGCTGCACTC AATGTACACC GACATGTGGA 5820

GTGAAGAGTA TCAGTGTGCA TGGCTGGATA TGTATCACCG CGTCTTTGAT CGCGTCAGCG 5880

CCGTCGTCGG TGAACAGGTA TGGAATTTCG CCGATTTTGC GACCTCGCAA GGCATATTGC 5940

GCGTTGGCGG TAACAAGAAA GGGATCTTCA CTCGCGACCG CAAACCGAAG TCGGCGGCTT 6000

TTCTGCTGCA AAAACGCTGG ACTGGCATGA ACTTCGGTGA AAAACCGCAG CAGGGAGGCA 6060

AACAATGAAT CAACAACTCT CCTGGCGCAC CATCGTCGGC TACAGCCTCG GTGGGGAATT 6120

GGAGCTCGAA TTTCCCCGAT CGTTCAAACA TTTGGCAATA AAGTTTCTTA AGATTGAATC 6180

CTGTTGCCGG TCTTGCGATG ATTATCATAT AATTTCTGTT GAATTACGTT AAGCATGTAA 6240

TAATTAACAT GTAATGCATG ACGTTATTTA TGAGATGGGT TTTTATGATT AGAGTCCCGC 6300

AATTATACAT TTAATACGCG ATAGAAAACA AAATATAGCG CGCAAACTAG GATAAATTAT 6360

CGCGCGCGGT GTCATCTATG TTACTAGATC GATCGGGAAT TAAGCTTATC GATACCGTCG 6420

ACCTCGAGGG GGGGCCCGGT ACCCAATTCG CCCTATAGTG AGTCGTATTA CAATTCACTG 6480

GCCGTCGTTT TACAACGTCG TGACTGGGAA AACCCTGGCG TTACCCAACT TAATCGCCTT 6540

GCAGCACATC CCCCTTTCGC CAGCTGGCGT AATAGCGAAG AGGCCCGCAC CGATCGCCCT 6600

TCCCAACAGT TGCGCAGCCT GAATGGCGAA TGGCGCGAAA TTGTAAACGT TAATATTTTG 6660

TTAAAATTCG CGTTAAATTT TTGTTAAATC AGCTCATTTT TTAACCAATA GGCCGAAATC 6720

GGCAAAATCC CTTATAAATC AAAAGAATAG ACCGAGATAG GGTTGAGTGT TGTTCCAGTT 6780

TGGAACAAGA GTCCACTATT AAAGAACGTG GACTCCAACG TCAAAGGGCG AAAAACCGTC 6840

TATCAGGGCG ATGGCCCACT ACGTGAACCA TCACCCTAAT CAAGTTTTTT GGGGTCGAGG 6900

TGCCGTAAAG CACTAAATCG GAACCCTAAA GGGAGCCCCC GATTTAGAGC TTGACGGGGA 6960

AAGCCGGCGA ACGTGGCGAG AAAGGAAGGG AAGAAAGCGA AAGGAGCGGG CGCTAGGGCG 7020

CTGGCAAGTG TAGCGGTCAC GCTGCGCGTA ACCACCACAC CCGCCGCGCT TAATGCGCCG 7080

CTACAGGGCG CGTCCCAGGT GGCACTTTTC GGGGAAATGT GCGCGGAACC CCTATTTGTT 7140

TATTTTTCTA AATACATTCA AATATGTATC CGCTCATGAG ACAATAACCC TGATAAATGC 7200

TTCAATAATA TTGAAAAAGG AAGAGTATGA GTATTCAACA TTTCCGTGTC GCCCTTATTC 7260

CCTTTTTTGC GGCATTTTGC CTTCCTGTTT TTGCTCACCC AGAAACGCTG GTGAAAGTAA 7320

AAGATGCTGA AGATCAGTTG GGTGCACGAG TGGGTTACAT CGAACTGGAT CTCAACAGCG 7380

GTAAGATCCT TGAGAGTTTT CGCCCCGAAG AACGTTTTCC AATGATGAGC ACTTTTAAAG 7440

TTCTGCTATG TGGCGCGGTA TTATCCCGTA TTGACGCCGG GCAAGAGCAA CTCGGTCGCC 7500

GCATACACTA TTCTCAGAAT GACTTGGTTG AGTACTCACC AGTCACAGAA AAGCATCTTA 7560

CGGATGGCAT GACAGTAAGA GAATTATGCA GTGCTGCCAT AACCATGAGT GATAACACTG 7620

CGGCCAACTT ACTTCTGACA ACGATCGGAG GACCGAAGGA GCTAACCGCT TTTTTGCACA 7680

ACATGGGGGA TCATGTAACT CGCCTTGATC GTTGGGAACC GGAGCTGAAT GAAGCCATAC 7740

CAAACGACGA GCGTGACACC ACGATGCCTG TAGCAATGGC AACAACGTTG CGCAAACTAT 7800

TAACTGGCGA ACTACTTACT CTAGCTTCCC GGCAACAATT AATAGACTGG ATGGAGGCGG 7860

ATAAAGTTGC AGGACCACTT CTGCGCTCGG CCCTTCCGGC TGGCTGGTTT ATTGCTGATA 7920

AATCTGGAGC CGGTGAGCGT GGGTCTCGCG GTATCATTGC AGCACTGGGG CCAGATGGTA 7980

AGCCCTCCCG TATCGTAGTT ATCTACACGA CGGGGAGTCA GGCAACTATG GATGAACGAA 8040

ATAGACAGAT CGCTGAGATA GGTGCCTCAC TGATTAAGCA TTGGTAACTG TCAGACCAAG 8100

TTTACTCATA TATACTTTAG ATTGATTTAA AACTTCATTT TTAATTTAAA AGGATCTAGG 8160

TGAAGATCCT TTTTGATAAT CTCATGACCA AAATCCCTTA ACGTGAGTTT TCGTTCCACT 8220

GAGCGTCAGA CCCCGTAGAA AAGATCAAAG GATCTTCTTG AGATCCTTTT TTTCTGCGCG 8280

TAATCTGCTG CTTGCAAACA AAAAAACCAC CGCTACCAGC GGTGGTTTGT TTGCCGGATC 8340

AAGAGCTACC AACTCTTTTT CCGAAGGTAA CTGGCTTCAG CAGAGCGCAG ATACCAAATA 8400

CTGTCCTTCT AGTGTAGCCG TAGTTAGGCC ACCACTTCAA GAACTCTGTA GCACCGCCTA 8460

CATACCTCGC TCTGCTAATC CTGTTACCAG TGGCTGCTGC CAGTGGCGAT AAGTCGTGTC 8520

TTACCGGGTT GGACTCAAGA CGATAGTTAC CGGATAAGGC GCAGCGGTCG GGCTGAACGG 8580

GGGGTTCGTG CACACAGCCC AGCTTGGAGC GAACGACCTA CACCGAACTG AGATACCTAC 8640

AGCGTGAGCT ATGAGAAAGC GCCACGCTTC CCGAAGGGAG AAAGGCGGAC AGGTATCCGG 8700

TAAGCGGCAG GGTCGGAACA GGAGAGCGCA CGAGGGAGCT TCCAGGGGGA AACGCCTGGT 8760

ATCTTTATAG TCCTGTCGGG TTTCGCCACC TCTGACTTGA GCGTCGATTT TTGTGATGCT 8820

CGTCAGGGGG GCGGAGCCTA TGGAAAAACG CCAGCAACGC GGCCTTTTTA CGGTTCCTGG 8880

CCTTTTGCTG GCCTTTTGCT CACATGTTCT TTCCTGCGTT ATCCCCTGAT TCTGTGGATA 8940

ACCGTATTAC CGCCTTTGAG TGAGCTGATA CCGCTCGCCG CAGCCGAACG ACCGAGCGCA 9000

GCGAGTCAGT GAGCGAGGAA GCGGAAGAGC GCCCAATACG CAAACCGCCT CTCCCCGCGC 9060

GTTGGCCGAT TCATTAATGC AGCTGGCACG ACAGGTTTCC CGACTGGAAA GCGGGCAGTG 9120

AGCGCAACGC AATTAATGTG AGTTAGCTCA CTCATTAGGC ACCCCAGGCT TTACACTTTA 9180

TGCTTCCGGC TCGTATGTTG TGTGGAATTG TGAGCGGATA ACAATTTCAC ACAGGAAACA 9240

GCTATGACCA TGATTACGCC AAGCTCGGAA TTAACCCTCA CTAAAGGGAA CAAAAGCTG 9299

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9408 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GAGCTCCACC GCGGTGGCGG CCGCTCTAGA ACTAGTGGAT CCTCTAGAGT CGACCATGGC 60

CAGTTGCCGG TGGAGCAGGT AAAAACACCG TAGCGTAGCA GCCAGGCGGA AGCAGACGCA 120

CAGCACAGGT TGGTTATGAT AGTCAGCCGG GCCACATGTG TGTAGTTGGT ACACTGATAC 180

GCTTACACTG TCTCTCCTTT CTTTTTTATT TGTCACCTTT GGTCGAGCTT ACATAATTGT 240

GTGACTAAAA AAAGGTCACT TCATTCAGAA ATTTAGGGTT GTGGGAATTT TGGATTTTAT 300

TGTGTCTGTA TAGAGTAGCT ATAGCTAGCT AGCTAGATGT GATGTTAATA ATTATGACGA 360

TGAGATTGGC CCGCTTGGCC GCTTGCATTG TCTCCCTAGC TCAATAATGT TTTGAGTTTG 420

TCTTGCCTTT CTTTCAGCTC TAACAAATTG GAGTAGGGAT GACTGAGATA CATATATAAA 480

AGCGAAAACC GCTGCTCTCT GTTAATTATT GCACATCACA CATAGGCCAA GCCTTAAGGA 540

CAATCAACTA AGGATGGTAA TAACTAAGGC TAGTGAGGTC GAACTAGGGA TGTTAATATA 600

CTCTAGATTT TAGACTATAA AATTTAAGGA TCGAATCAGA TTAGTATCGA ACTATATTTA 660

TATTCATTTC TAAACTAAAT TAATTAAGCA CCCTAAATTA TTGTGATGAA GAGACATTTC 720

GATCGTGATC CATTATTACT CCTTGGTCAA ACTAATCTCG TTTTATGTCA CTATTTCATC 780

ATCTTTTTTG CGAACGGGTT TATAGCCCGT GTTCCATTAT GAGGACATGA ACGGTTTAAA 840

CAAAGTTACA TATCATCCCA GCTAGCTACC TAGATTGGAA GCATGGGTTC GGTATATATA 900

TATAGTTTAT ATATTTGGTA TATATATATA TATATATATA TATATATATA TATCACACGT 960

CAGCTTATAT TACGTAAAGT GGGGTTAGTT TTCAAGAAGC GTGGGACCAG TCACCTCTGC 1020

AGTCTGACCT TGGCTTCAGC TTCGACAGCA AACAGTCATC TCTTGGAAGC TAAGGACAGT 1080

CTCCAACAGT CAACAAAGCA GCGGTCTGCT TGTAGTTCTC CCTTGCACGA CCAGCTATAT 1140

CTAGCATCAT AACAACGGTA AGATCATCTC TAGCACGACA AACTTAGTTT AATTAATTAT 1200

GTCTAATCCG TTGTTGTTAG CTTAAACTTT CTAGCCTCCT ATGCTAAGAG AGTTCTCTAG 1260

TTCTACTCAG GTGGATTGAT ATATAAATTG GGAATCTTCT AGGCGTCACA AGGTATGGTA 1320

CACATCAATC AATGAACGGA CAAAGCAACG GTAAGATCCG ACCCAGTAAA AGTAATAGCG 1380

TTAGGGCATG TACAACCTAG ACACTGATGC ACAGTACTCC AAGTATAAGA CACAACTAAA 1440

ACACAACATA ATAATACAGT GGTTATATCT AAAACATGTG TCTTACCATA TTCATTGTAC 1500

CAATTAGAAC ATTTAATAAA TTAAAGTGAC CAATCAGCTA GCCTCCTGTC TCGAACATAG 1560

AGCTAAGACA TTGTGTCTTC GTCAAGATAC ATGTCTTAAG TTTTTTTATA TTCACTCCCA 1620

AAGACACACT CTAAGACACA ACGTAACACA CCCATTGTAC ATGCTCTTAA CCTAAGTTAT 1680

CATGGATGAC CACGCGTGGC AATTAAAAAA ATAATTTTTG CCTCCTAAAA CCTCTTTCTT 1740

AATTGGTTCT TGCTTGCAAA TCACCAGCGA ACCCATATGA AAGGATGCTC AAAATCTGGC 1800

CACCGCATCA GGGTTGGTGA ATGCAAVGTA AAAAATAATG CATAAATCAG CTCTCTGATC 1860

AGTTATATAA TCGTGCCTTT TAATTATTCA TGCCAGCTTT ATCTGACTCA CGAAATCATT 1920

GATAAATTAT TCCTCAGCTG TATTAGAAAG AGCAGTGTTG TTTAACTTGG AAAGTGATGT 1980

GGAAGCGTGT GATTGCGGTT GAGCTTGTAT AGGAGTAAAA TGAGGAACAG TAGGAAAATA 2040

ATTTTTTCGG ATTAAAACCG GTTGTTTGGA CTGCGGCAGA TACAATTCAT AGAGATAAAA 2100

ACACCGTAGA AGTATTAGAA GCCGATAAAG ATTAAACCCA AATGAACGAA CAGGCTAAAC 2160

AAATCCGGCG CCTCAAAAGT CAAGAGCAGG TACTGGGCTG TCTTGCACAC GTCGCTTTTT 2220

GTCTCCCCCT GGCCCCTGGG TGAGAGTAGT AGGGATGCTA AAGTTTGCTT TCTCTTTTTG 2280

AGGCATGTGA TAGGCTCTTG TTAGTTGCTA GGGCTATGTT TATAATATTT GCGCTTTTAC 2340

CTATGTACGT AAGAACCGGA TGGAATAATG CTATGCAGGA ACCAATTATG TTTGGTCGAA 2400

ATATATAGTG ACCTATCATA ATGTTATCCC TGTTCATGTA CCTAGGTGGC TAATGATATA 2460

CGGCATATGA ATACAGTAAT CATCCAAGCA CGTAAAAACT CGCTAGACGT TTATGCCTGC 2520

TAGCCTGCTG GGTGTGTAGA CTGGAGTACT GGACAAACAT CGCAATACAG AGGTACAGTA 2580

TTTGTCTAGA CAATGATATA CATAGATAAA AACCACTGTT GTAACTTGTA AGCCACTAGC 2640

TCACGTTCTC CATGAGCTCT TCTCTCTGCT GTTTCTTCCT CTGCTAACTG CGTTATGATA 2700

TGACGTCGTA TAAATAATCT CACAATACTT CCTTATTTTC AGCATGGCCT CTTTTATGTT 2760

TATTTAACAG TAGCAACCAA CGCCGCTCGA TGTTTCCTTC AAGAAACGGC CACTCACTAT 2820

GTGGTGTGCA GAAGAACAAA TGTAAGCAGC TCCTACAGGT ACCAGTAGTC ATGTCAGTGT 2880

GGAAGCTTTC CAACCAACGC CTCCTTCGAG GAACCTGGTC GTGCTGACAT GAATGTAGGC 2940

CATGCAAGCA CAAGCACCTA ACGCGAATCA TCACGACGCG CCGTGTACTG GGCGTTGGTA 3000

CATCACACCC CGCGTTTGAC CTGATCGGAA GCATGCGTGT GTGTTGGCTG CAGGACCGGC 3060

TATAGGTTTC CTGCATTGGA CAGCAGAAGC CAGTCATGTT AGGCACTCAC GCGCTCCTGC 3120

CGTTTGATGA ATCATCCGGT CTTTCGTATT GATCACTAGT TCACTACGCT GATATAGCAA 3180

ATTTTAAGAT GTGAAACCAC GAGACGAGCG ATAAATCTTA GACGTTACCT ATCCATATGA 3240

AGCTTGTGCG AAAAAAAGGC GTGCCGCTGT AGCATCATTC GTATACACTT TTGTCCCCAA 3300

AGACAGGGAT ACGAATCCAT GCTCGACAGA ACCCTCCCTT CCCTGCAGAT AACGACACTT 3360

AAGTATAACA AAAGTAGTTG GATTATTTCA GAAGCAAAAT CTCACTTTTC GCTGGCCTTT 3420

TTGTACTTTG GTTACTTGAG TTCAGACAGT GTATGCTATA TTGTCATGTG CTGCGTAAGG 3480

TTTAAATATG GTTCGACAAA TATATCAGTA TATCACTACT TTGTTATGGG TGGGGCCTAG 3540

CACAAACTTG ATACAGCTAG GATAAAGTTA GAACGATGAC TGATCTACTG TAAAGCGACA 3600

CCTGTCCTGT TATGGTAGTT TAAGTCCATT CCTGGACGAC TCCAGATCCA GGATATGATG 3660

CTGTTACATA ATGCGATTGT TCACAATAAA ATTGCATGAT GTTCTTCTAC TCTTTAGGCA 3720

GTTTTGTTCA ACAGGCAAGT TGCATAATGC ATGTGCATAT ATGAGCAGCA TAATCATCAA 3780

TTAATCATAG GTTCGTCATT TTAGTTTCAC TCCTTCACAT TATTCCAGCC CTTGAAGAAA 3840

AATGTAGCAG TGCTTGCTGT TTAATAAGTG GCAGAGCTGT TTTCACTCCA CCTACGCTTG 3900

TCTAGGACCA AAATTTTAAT CTGTCACTTT GAGCTAAAAC TGAAGCACCA AACCGCTACA 3960

AAAGAACGTA GGAGCTGAAT TGTAACTTGA TGGGATTACT ATAGCAGTTG CTACAGTTCT 4020

AGCTAGCTAC CTTATTCTAT ACGCATCACC CTAACAACCC GGCTGACTGC TGCATCTGAC 4080

CCCACCGTCC CCTGCTCCAA ACCAACTCTC CTTTCCTTGC ATGCACTACA CCCACTTCCT 4140

GCAGCTATAT ATACCACCAT ATGCCCATCT TATGAAACCA TCCACAAGAG GAGAAGAAAC 4200

AATCAACCAG CAACACTCTT CTCTTATAAC ATAGTACAGC GAAGGTAACT CACATGGCAA 4260

CTTCCATGGT CCGTCCTGTA GAAACCCCAA CCCGTGAAAT CAAAAAACTC GACGGCCTGT 4320

GGGCATTCAG TCTGGATCGC GAAAACTGTG GAATTGATCA GCGTTGGTGG GAAAGCGCGT 4380

TACAAGAAAG CCGGGCAATT GCTGTGCCAG GCAGTTTTAA CGATCAGTTC GCCGATGCAG 4440

ATATTCGTAA TTATGCGGGC AACGTCTGGT ATCAGCGCGA AGTCTTTATA CCGAAAGGTT 4500

GGGCAGGCCA GCGTATCGTG CTGCGTTTCG ATGCGGTCAC TCATTACGGC AAAGTGTGGG 4560

TCAATAATCA GGAAGTGATG GAGCATCAGG GCGGCTATAC GCCATTTGAA GCCGATGTCA 4620

CGCCGTATGT TATTGCCGGG AAAAGTGTAC GTATCACCGT TTGTGTGAAC AACGAACTGA 4680

ACTGGCAGAC TATCCCGCCG GGAATGGTGA TTACCGACGA AAACGGCAAG AAAAAGCAGT 4740

CTTACTTCCA TGATTTCTTT AACTATGCCG GAATCCATCG CAGCGTAATG CTCTACACCA 4800

CGCCGAACAC CTGGGTGGAC GATATCACCG TGGTGACGCA TGTCGCGCAA GACTGTAACC 4860

ACGCGTCTGT TGACTGGCAG GTGGTGGCCA ATGGTGATGT CAGCGTTGAA CTGCGTGATG 4920

CGGATCAACA GGTGGTTGCA ACTGGACAAG GCACTAGCGG GACTTTGCAA GTGGTGAATC 4980

CGCACCTCTG GCAACCGGGT GAAGGTTATC TCTATGAACT GTGCGTCACA GCCAAAAGCC 5040

AGACAGAGTG TGATATCTAC CCGCTTCGCG TCGGCATCCG GTCAGTGGCA GTGAAGGGCG 5100

AACAGTTCCT GATTAACCAC AAACCGTTCT ACTTTACTGG CTTTGGTCGT CATGAAGATG 5160

CGGACTTACG TGGCAAAGGA TTCGATAACG TGCTGATGGT GCACGACCAC GCATTAATGG 5220

ACTGGATTGG GGCCAACTCC TACCGTACCT CGCATTACCC TTACGCTGAA GAGATGCTCG 5280

ACTGGGCAGA TGAACATGGC ATCGTGGTGA TTGATGAAAC TGCTGCTGTC GGCTTTAACC 5340

TCTCTTTAGG CATTGGTTTC GAAGCGGGCA ACAAGCCGAA AGAACTGTAC AGCGAAGAGG 5400

CAGTCAACGG GGAAACTCAG CAAGCGCACT TACAGGCGAT TAAAGAGCTG ATAGCGCGTG 5460

ACAAAAACCA CCCAAGCGTG GTGATGTGGA GTATTGCCAA CGAACCGGAT ACCCGTCCGC 5520

AAGTGCACGG GAATATTTCG CCACTGGCGG AAGCAACGCG TAAACTCGAC CCGACGCGTC 5580

CGATCACCTG CGTCAATGTA ATGTTCTGCG ACGCTCACAC CGATACCATC AGCGATCTCT 5640

TTGATGTGCT GTGCCTGAAC CGTTATTACG GATGGTATGT CCAAAGCGGC GATTTGGAAA 5700

CGGCAGAGAA GGTACTGGAA AAAGAACTTC TGGCCTGGCA GGAGAAACTG CATCAGCCGA 5760

TTATCATCAC CGAATACGGC GTGGATACGT TAGCCGGGCT GCACTCAATG TACACCGACA 5820

TGTGGAGTGA AGAGTATCAG TGTGCATGGC TGGATATGTA TCACCGCGTC TTTGATCGCG 5880

TCAGCGCCGT CGTCGGTGAA CAGGTATGGA ATTTCGCCGA TTTTGCGACC TCGCAAGGCA 5940

TATTGCGCGT TGGCGGTAAC AAGAAAGGGA TCTTCACTCG CGACCGCAAA CCGAAGTCGG 6000

CGGCTTTTCT GCTGCAAAAA CGCTGGACTG GCATGAACTT CGGTGAAAAA CCGCAGCAGG 6060

GAGGCAAACA ATGAATCAAC AACTCTCCTG GCGCACCATC GTCGGCTACA GCCTCGGGAA 6120

TTGCTACCGA GCTTCTCGAG GGCACTGAAG TCGCTTGATG TGCTGAATTG TTTGTGATGT 6180

TGGTGGCGTA TTTTGTTTAA ATAAGTAAGC ATGGCTGTGA TTTTATCATA TGATCGATCT 6240

TTGGGGTTTT ATTTAACACA TTGTAAAATG TGTATCTATT AATAACTCAA TGTATAAGAT 6300

GTGTTCATTC TTCGGTTGCC ATAGATCTGC TTATTTGACC TGTGATGTTT TGACTCCAAA 6360

AACCAAAATC ACAACTCAAT AAACTCATGG AATATGTCCA CCTGTTTCTT GAAGAGTTCA 6420

TCTACCATTC CAGTTGGCAT TTATCAGTGT TGCAGCGGCG CTGTGCTTTG TAACATAACA 6480

ATTGTTCACG GCATATATCC AAATCTAGAG AAGCTTATCG ATACCGTCGA CCTCGAGGGG 6540

GGGCCCGGTA CCCAATTCGC CCTATAGTGA GTCGTATTAC AATTCACTGG CCGTCGTTTT 6600

ACAACGTCGT GACTGGGAAA ACCCTGGCGT TACCCAACTT AATCGCCTTG CAGCACATCC 6660

CCCTTTCGCC AGCTGGCGTA ATAGCGAAGA GGCCCGCACC GATCGCCCTT CCCAACAGTT 6720

GCGCAGCCTG AATGGCGAAT GGCGCGAAAT TGTAAACGTT AATATTTTGT TAAAATTCGC 6780

GTTAAATTTT TGTTAAATCA GCTCATTTTT TAACCAATAG GCCGAAATCG GCAAAATCCC 6840

TTATAAATCA AAAGAATAGA CCGAGATAGG GTTGAGTGTT GTTCCAGTTT GGAACAAGAG 6900

TCCACTATTA AAGAACGTGG ACTCCAACGT CAAAGGGCGA AAAACCGTCT ATCAGGGCGA 6960

TGGCCCACTA CGTGAACCAT CACCCTAATC AAGTTTTTTG GGGTCGAGGT GCCGTAAAGC 7020

ACTAAATCGG AACCCTAAAG GGAGCCCCCG ATTTAGAGCT TGACGGGGAA AGCCGGCGAA 7080

CGTGGCGAGA AAGGAAGGGA AGAAAGCGAA AGGAGCGGGC GCTAGGGCGC TGGCAAGTGT 7140

AGCGGTCACG CTGCGCGTAA CCACCACACC CGCCGCGCTT AATGCGCCGC TACAGGGCGC 7200

GTCCCAGGTG GCACTTTTCG GGGAAATGTG CGCGGAACCC CTATTTGTTT ATTTTTCTAA 7260

ATACATTCAA ATATGTATCC GCTCATGAGA CAATAACCCT GATAAATGCT TCAATAATAT 7320

TGAAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTCG CCCTTATTCC CTTTTTTGCG 7380

GCATTTTGCC TTCCTGTTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA AGATGCTGAA 7440

GATCAGTTGG GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG TAAGATCCTT 7500

GAGAGTTTTC GCCCCGAAGA ACGTTTTCCA ATGATGAGCA CTTTTAAAGT TCTGCTATGT 7560

GGCGCGGTAT TATCCCGTAT TGACGCCGGG CAAGAGCAAC TCGGTCGCCG CATACACTAT 7620

TCTCAGAATG ACTTGGTTGA GTACTCACCA GTCACAGAAA AGCATCTTAC GGATGGCATG 7680

ACAGTAAGAG AATTATGCAG TGCTGCCATA ACCATGAGTG ATAACACTGC GGCCAACTTA 7740

CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTTGCACAA CATGGGGGAT 7800

CATGTAACTC GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC AAACGACGAG 7860

CGTGACACCA CGATGCCTGT AGCAATGGCA ACAACGTTGC GCAAACTATT AACTGGCGAA 7920

CTACTTACTC TAGCTTCCCG GCAACAATTA ATAGACTGGA TGGAGGCGGA TAAAGTTGCA 7980

GGACCACTTC TGCGCTCGGC CCTTCCGGCT GGCTGGTTTA TTGCTGATAA ATCTGGAGCC 8040

GGTGAGCGTG GGTCTCGCGG TATCATTGCA GCACTGGGGC CAGATGGTAA GCCCTCCCGT 8100

ATCGTAGTTA TCTACACGAC GGGGAGTCAG GCAACTATGG ATGAACGAAA TAGACAGATC 8160

GCTGAGATAG GTGCCTCACT GATTAAGCAT TGGTAACTGT CAGACCAAGT TTACTCATAT 8220

ATACTTTAGA TTGATTTAAA ACTTCATTTT TAATTTAAAA GGATCTAGGT GAAGATCCTT 8280

TTTGATAATC TCATGACCAA AATCCCTTAA CGTGAGTTTT CGTTCCACTG AGCGTCAGAC 8340

CCCGTAGAAA AGATCAAAGG ATCTTCTTGA GATCCTTTTT TTCTGCGCGT AATCTGCTGC 8400

TTGCAAACAA AAAAACCACC GCTACCAGCG GTGGTTTGTT TGCCGGATCA AGAGCTACCA 8460

ACTCTTTTTC CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAAATAC TGTCCTTCTA 8520

GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCCTAC ATACCTCGCT 8580

CTGCTAATCC TGTTACCAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT TACCGGGTTG 8640

GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCGG GCTGAACGGG GGGTTCGTGC 8700

ACACAGCCCA GCTTGGAGCG AACGACCTAC ACCGAACTGA GATACCTACA GCGTGAGCTA 8760

TGAGAAAGCG CCACGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCGGT AAGCGGCAGG 8820

GTCGGAACAG GAGAGCGCAC GAGGGAGCTT CCAGGGGGAA ACGCCTGGTA TCTTTATAGT 8880

CCTGTCGGGT TTCGCCACCT CTGACTTGAG CGTCGATTTT TGTGATGCTC GTCAGGGGGG 8940

CGGAGCCTAT GGAAAAACGC CAGCAACGCG GCCTTTTTAC GGTTCCTGGC CTTTTGCTGG 9000

CCTTTTGCTC ACATGTTCTT TCCTGCGTTA TCCCCTGATT CTGTGGATAA CCGTATTACC 9060

GCCTTTGAGT GAGCTGATAC CGCTCGCCGC AGCCGAACGA CCGAGCGCAG CGAGTCAGTG 9120

AGCGAGGAAG CGGAAGAGCG CCCAATACGC AAACCGCCTC TCCCCGCGCG TTGGCCGATT 9180

CATTAATGCA GCTGGCACGA CAGGTTTCCC GACTGGAAAG CGGGCAGTGA GCGCAACGCA 9240

ATTAATGTGA GTTAGCTCAC TCATTAGGCA CCCCAGGCTT TACACTTTAT GCTTCCGGCT 9300

CGTATGTTGT GTGGAATTGT GAGCGGATAA CAATTTCACA CAGGAAACAG CTATGACCAT 9360

GATTACGCCA AGCTCGGAAT TAACCCTCAC TAAAGGGAAC AAAAGCTG 9408

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TTATCTCGAG GGCACTGAAG TCGCTTGATG TGCTGAATT 39

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GGGGAAGCTT CTCTAGATTT GGATATATGC CGTGAACAAT TG 42

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9335 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

AGCTTGCATG CCTGCAGGCC GGCCTTAATT AAGCGGCCGC CAGTGTGATG GATATCTGCA 60

GAATTCGGCT TGGGGGATCC TCTAGACAAT GATATACATA GATAAAAACC ACTGTTGTAA 120

CTTGTAAGCC ACTAGCTCAC GTTCTCCATG AGCTCTTCTC TCTGCTGTTT CTTCCTCTGC 180

TAACTGCGTT ATGATATGAC GTCGTATAAA TAATCTCACA ATACTTCCTT ATTTTCAGCA 240

TGGCCTCTTT TATGTTTATT TAACAGTAGC AACCAACGCC GCTCGATGTT TCCTTCAAGA 300

AACGGCCACT CACTATGTGG TGTGCAGAAG AACAAATGTA AGCAGCTCCT ACAGGTACCA 360

GTAGTCATGT CAGTGTGGAA GCTTTCCAAC CAACGCCTCC TTCGAGGAAC CTGGTCGTGC 420

TGACATGAAT GTAGGCCATG CAAGCACAAG CACCTAACGC GAATCATCAC GACGCGCCGT 480

GTACTGGGCG TTGGTACATC ACACCCCGCG TTTGACCTGA TCGGAAGCAT GCGTGTGTGT 540

TGGCTGCAGG ACCGGCTATA GGTTTCCTGC ATTGGACAGC AGAAGCCAGT CATGTTAGGC 600

ACTCACGCGC TCCTGCCGTT TGATGAATCA TCCGGTCTTT CGTATTGATC ACTAGTTCAC 660

TACGCTGATA TAGCAAATTT TAAGATGTGA AACCACGAGA CGAGCGATAA ATCTTAGACG 720

TTACCTATCC ATATGAAGCT TGTGCGAAAA AAAGGCGTGC CGCTGTAGCA TCATTCGTAT 780

ACACTTTTGT CCCCAAAGAC AGGGATACGA ATCCATGCTC GACAGAACCC TCCCTTCCCT 840

GCAGATAACG ACACTTAAGT ATAACAAAAG TAGTTGGATT ATTTCAGAAG CAAAATCTCA 900

CTTTTCGCTG GCCTTTTTGT ACTTTGGTTA CTTGAGTTCA GACAGTGTAT GCTATATTGT 960

CATGTGCTGC GTAAGGTTTA AATATGGTTC GACAAATATA TCAGTATATC ACTACTTTGT 1020

TATGGGTGGG GCCTAGCACA AACTTGATAC AGCTAGGATA AAGTTAGAAC GATGACTGAT 1080

CTACTGTAAA GCGACACCTG TCCTGTTATG GTAGTTTAAG TCCATTCCTG GACGACTCCA 1140

GATCCAGGAT ATGATGCTGT TACATAATGC GATTGTTCAC AATAAAATTG CATGATGTTC 1200

TTCTACTCTT TAGGCAGTTT TGTTCAACAG GCAAGTTGCA TAATGCATGT GCATATATGA 1260

GCAGCATAAT CATCAATTAA TCATAGGTTC GTCATTTTAG TTTCACTCCT TCACATTATT 1320

CCAGCCCTTG AAGAAAAATG TAGCAGTGCT TGCTGTTTAA TAAGTGGCAG AGCTGTTTTC 1380

ACTCCACCTA CGCTTGTCTA GGACCAAAAT TTTAATCTGT CACTTTGAGC TAAAACTGAA 1440

GCACCAAACC GCTACAAAAG AACGTAGGAG CTGAATTGTA ACTTGATGGG ATTACTATAG 1500

CAGTTGCTAC AGTTCTAGCT AGCTACCTTA TTCTATACGC ATCACCCTAA CAACCCGGCT 1560

GACTGCTGCA TCTGACCCCA CCGTCCCCTG CTCCAAACCA ACTCTCCTTT CCTTGCATGC 1620

ACTACACCCA CTTCCTGCAG CTATATATAC CACCATATGC CCATCTTATG AAACCATCCA 1680

CAAGAGGAGA AGAAACAATC AACCAGCAAC ACTCTTCTCT TATAACATAG TACAGCGAAG 1740

GAGATCCTGA CTGCTTTGTC AAGGTTCAAT TCTGCTTCCT CTGTTATGTT CTTTATATTA 1800

CATGCTCTGA CAAAGCTATA AAGCTTGATA CTGCAGTATA ATATAACAAG TTAGCTACAC 1860

AAGTTTTGTA CTTCAAGTCT TTTAACTATA TGTTGGTGCA ATAAGATTAT GAGTAATCCA 1920

TATGAAGGTG TTGCAAGAGA ACATGAAAGG CAAAGATAAA CGGATGAACC CATTACTAGC 1980

TTTGGCTGTA TCAGACCAAT AACTTGAAAT GCACTTGTGC TAGCATGCCT AAGTATTAGA 2040

AAAGGTAGCA TGGGAGAATC TATATTATTT TGGCTAACTT CTTTAGTTAC TATTGATTGA 2100

TGAGAAAGCC TACCATTGCC CATGCCAGCC CTAATGTCCC GGTGACATGA TTGAGCCAGT 2160

ACTATGATTA ATTTACTCTA TTGTTCTCCT TTTTTGAGTG CTGTATAAGA TGTCCTTTTT 2220

TTGAGCCACT CGAGAAGATG TTTACTTAAC TCTAGTGCGC AATGATTGGA GCTCTCAGTG 2280

CAACGCATGT GCTCTGTAAT CTACTGTCAC CACTACTCTG TAGTGTGTGC TTAAACTCTA 2340

AACTATTCCA CGTGGCTAGT AATTACCAAT CATTTACAAC ACTGTTACAT GTGTAGGGCT 2400

GCGATCCATG GTCCGTCCTG TAGAAACCCC AACCCGTGAA ATCAAAAAAC TCGACGGCCT 2460

GTGGGCATTC AGTCTGGATC GCGAAAACTG TGGAATTGAT CAGCGTTGGT GGGAAAGCGC 2520

GTTACAAGAA AGCCGGGCAA TTGCTGTGCC AGGCAGTTTT AACGATCAGT TCGCCGATGC 2580

AGATATTCGT AATTATGCGG GCAACGTCTG GTATCAGCGC GAAGTCTTTA TACCGAAAGG 2640

TTGGGCAGGC CAGCGTATCG TGCTGCGTTT CGATGCGGTC ACTCATTACG GCAAAGTGTG 2700

GGTCAATAAT CAGGAAGTGA TGGAGCATCA GGGCGGCTAT ACGCCATTTG AAGCCGATGT 2760

CACGCCGTAT GTTATTGCCG GGAAAAGTGT ACGTATCACC GTTTGTGTGA ACAACGAACT 2820

GAACTGGCAG ACTATCCCGC CGGGAATGGT GATTACCGAC GAAAACGGCA AGAAAAAGCA 2880

GTCTTACTTC CATGATTTCT TTAACTATGC CGGAATCCAT CGCAGCGTAA TGCTCTACAC 2940

CACGCCGAAC ACCTGGGTGG ACGATATCAC CGTGGTGACG CATGTCGCGC AAGACTGTAA 3000

CCACGCGTCT GTTGACTGGC AGGTGGTGGC CAATGGTGAT GTCAGCGTTG AACTGCGTGA 3060

TGCGGATCAA CAGGTGGTTG CAACTGGACA AGGCACTAGC GGGACTTTGC AAGTGGTGAA 3120

TCCGCACCTC TGGCAACCGG GTGAAGGTTA TCTCTATGAA CTGTGCGTCA CAGCCAAAAG 3180

CCAGACAGAG TGTGATATCT ACCCGCTTCG CGTCGGCATC CGGTCAGTGG CAGTGAAGGG 3240

CGAACAGTTC CTGATTAACC ACAAACCGTT CTACTTTACT GGCTTTGGTC GTCATGAAGA 3300

TGCGGACTTA CGTGGCAAAG GATTCGATAA CGTGCTGATG GTGCACGACC ACGCATTAAT 3360

GGACTGGATT GGGGCCAACT CCTACCGTAC CTCGCATTAC CCTTACGCTG AAGAGATGCT 3420

CGACTGGGCA GATGAACATG GCATCGTGGT GATTGATGAA ACTGCTGCTG TCGGCTTTAA 3480

CCTCTCTTTA GGCATTGGTT TCGAAGCGGG CAACAAGCCG AAAGAACTGT ACAGCGAAGA 3540

GGCAGTCAAC GGGGAAACTC AGCAAGCGCA CTTACAGGCG ATTAAAGAGC TGATAGCGCG 3600

TGACAAAAAC CACCCAAGCG TGGTGATGTG GAGTATTGCC AACGAACCGG ATACCCGTCC 3660

GCAAGTGCAC GGGAATATTT CGCCACTGGC GGAAGCAACG CGTAAACTCG ACCCGACGCG 3720

TCCGATCACC TGCGTCAATG TAATGTTCTG CGACGCTCAC ACCGATACCA TCAGCGATCT 3780

CTTTGATGTG CTGTGCCTGA ACCGTTATTA CGGATGGTAT GTCCAAAGCG GCGATTTGGA 3840

AACGGCAGAG AAGGTACTGG AAAAAGAACT TCTGGCCTGG CAGGAGAAAC TGCATCAGCC 3900

GATTATCATC ACCGAATACG GCGTGGATAC GTTAGCCGGG CTGCACTCAA TGTACACCGA 3960

CATGTGGAGT GAAGAGTATC AGTGTGCATG GCTGGATATG TATCACCGCG TCTTTGATCG 4020

CGTCAGCGCC GTCGTCGGTG AACAGGTATG GAATTTCGCC GATTTTGCGA CCTCGCAAGG 4080

CATATTGCGC GTTGGCGGTA ACAAGAAAGG GATCTTCACT CGCGACCGCA AACCGAAGTC 4140

GGCGGCTTTT CTGCTGCAAA AACGCTGGAC TGGCATGAAC TTCGGTGAAA AACCGCAGCA 4200

GGGAGGCAAA CAATGAATCA ACAACTCTCC TGGCGCACCA TCGTCGGCTA CAGCCTCGGG 4260

AATTGCTACC GAGCTTCTCG AGGGCACTGA AGTCGCTTGA TGTGCTGAAT TGTTTGTGAT 4320

GTTGGTGGCG TATTTTGTTT AAATAAGTAA GCATGGCTGT GATTTTATCA TATGATCGAT 4380

CTTTGGGGTT TTATTTAACA CATTGTAAAA TGTGTATCTA TTAATAACTC AATGTATAAG 4440

ATGTGTTCAT TCTTCGGTTG CCATAGATCT GCTTATTTGA CCTGTGATGT TTTGACTCCA 4500

AAAACCAAAA TCACAACTCA ATAAACTCAT GGAATATGTC CACCTGTTTC TTGAAGAGTT 4560

CATCTACCAT TCCAGTTGGC ATTTATCAGT GTTGCAGCGG CGCTGTGCTT TGTAACATAA 4620

CAATTGTTCA CGGCATATAT CCAAATCTAG AGAAGCTTAT CGATACCGTC GACCTCGAGG 4680

GGGGGCCCGG TACCCAATTC GCCCTATAGT GAGTCGTATT ACAATTCACT GGCCGTCGTT 4740

TTACAACGTC GTGACTGGGA AAACCCTGGC GTTACCCAAC TTAATCGCCT TGCAGCACAT 4800

CCCCCTTTCG CCAGAAACGC CCGGGCATTT AAATGGCGCG CCGCGATCGC TTGCAGATCT 4860

GCATGGGTGG AGACTTTTCA ACAAAGGGTA ATATCCGGAA ACCTCCTCGG ATTCCATTGC 4920

CCAGCTATCT GTCACTTTAT TGTGAAGATA GTGGAAAAGG AAGGTGGCTC CTACAAATGC 4980

CATCATTGCG ATAAAGGAAA GGCCATCGTT GAAGATGCCT CTGCCGACAG TGGTCCCAAA 5040

GATGGACCCC CACCCACGAG GAGCATCGTG GAAAAAGAAG ACGTTCCAAC CACGTCTTCA 5100

AAGCAAGTGG ATTGATGTGA TCATCGATGG AGACTTTTCA ACAAAGGGTA ATATCCGGAA 5160

ACCTCCTCGG ATTCCATTGC CCAGCTATCT GTCACTTTAT TGTGAAGATA GTGGAAAAGG 5220

AAGGTGGCTC CTACAAATGC CATCATTGCG ATAAAGGAAA GGCCATCGTT GAAGATGCCT 5280

CTGCCGACAG TGGTCCCAAA GATGGACCCC CACCCACGAG GAGCATCGTG GAAAAAGAAG 5340

ACGTTCCAAC CACGTCTTCA AAGCAAGTGG ATTGATGTGA TATCTCCACT GACGTAAGGG 5400

ATGACGCACA ATCCCACTAT CCTTCGCAAG ACCCTTCCTC TATATAAGGA AGTTCATTTC 5460

ATTTGGAGAG AACACGGGGG ACTCTAGAGG ATCCAGCTGA AGGCTCGACA AGGCAGTCCA 5520

CGGAGGAGCT GATATTTGGT GGACAAGCTG TGGATAGGAG CAACCCTATC CCTAATATAC 5580

CAGCACCACC AAGTCAGGGC AATCCCCAGA TCAAGTGCAA AGGTCCGCCT TGTTTCTCCT 5640

CTGTCTCTTG ATCTGACTAA TCTTGGTTTA TGATTCGTTG AGTAATTTTG GGGAAAGCTC 5700

CTTTGCTGCT CCACACATGT CCATTCGAAT TTTACCGTGT TTAGCAAGGG CGAAAAGTTT 5760

GCATCTTGAT GATTTAGCTT GACTATGCGA TTGCTTTCCT GGACCCGTGC AGCTGCGGAC 5820

GGATCTGGGG CCATTTGTTC CAGGCACGGG ATAAGCATTC AGCCATGGCC CCAGAACGAC 5880

GCCCGGCCGA CATCCGCCGT GCCACCGAGG CGGACATGCC GGCGGTCTGC ACCATCGTCA 5940

ACCACTACAT CGAGACAAGC ACGGTCAACT TCCGTACCGA GCCGCAGGAA CCGCAGGAGT 6000

GGACGGACGA CCTCGTCCGT CTGCGGGAGC GCTATCCCTG GCTCGTCGCC GAGGTGGACG 6060

GCGAGGTCGC CGGCATCGCC TACGCGGGCC CCTGGAAGGC ACGCAACGCC TACGACTGGA 6120

CGGCCGAGTC GACCGTGTAC GTCTCCCCCC GCCACCAGCG GACGGGACTG GGCTCCACGC 6180

TCTACACCCA CCTGCTGAAG TCCCTGGAGG CACAGGGCTT CAAGAGCGTG GTCGCTGTCA 6240

TCGGGCTGCC CAACGACCCG AGCGTGCGCA TGCACGAGGC GCTCGGATAT GCCCCCCGCG 6300

GCATGCTGCG GGCGGCCGGC TTCAAGCACG GGAACTGGCA TGACGTGGGT TTCTGGCAGC 6360

TGGACTTCAG CCTGCCGGTA CCGCCCCGTC CGGTCCTGCC CGTCACCGAA ATCTGATGAG 6420

ATCTGAGCTC GAATTTCCCC GATCGTTCAA ACATTTGGCA ATAAAGTTTC TTAAGATTGA 6480

ATCCTGTTGC CGGTCTTGCG ATGATTATCA TATAATTTCT GTTGAATTAC GTTAAGCATG 6540

TAATAATTAA CATGTAATGC ATGACGTTAT TTATGAGATG GGTTTTTATG ATTAGAGTCC 6600

CGCAATTATA CATTTAATAC GCGATAGAAA ACAAAATATA GCGCGCAAAC TAGGATAAAT 6660

TATCGCGCGC GGTGTCATCT ATGTTACTAG ATCGATCGGG AATTCACTGG CCGTCGTTTT 6720

ACAACGTCGT GACTGGGAAA ACCCTGGCGT TACCCAACTT AATCGCCTTG CAGCACATCC 6780

CCCTTTCGCC AGCTGGCGTA ATAGCGAAGA GGCCCGCACC GATCGCCCTT CCCAACAGTT 6840

GCGCAGCCTG AATGGCGAAT GGCGCCTGAT GCGGTATTTT CTCCTTACGC ATCTGTGCGG 6900

TATTTCACAC CGCATATGGT GCACTCTCAG TACAATCTGC TCTGATGCCG CATAGTTAAG 6960

CCAGCCCCGA CACCCGCCAA CACCCGCTGA CGCGCCCTGA CGGGCTTGTC TGCTCCCGGC 7020

ATCCGCTTAC AGACAAGCTG TGACCGTCTC CGGGAGCTGC ATGTGTCAGA GGTTTTCACC 7080

GTCATCACCG AAACGCGCGA GACGAAAGGG CCTCGTGATA CGCCTATTTT TATAGGTTAA 7140

TGTCATGATA ATAATGGTTT CTTAGACGTC AGGTGGCACT TTTCGGGGAA ATGTGCGCGG 7200

AACCCCTATT TGTTTATTTT TCTAAATACA TTCAAATATG TATCCGCTCA TGAGACAATA 7260

ACCCTGATAA ATGCTTCAAT AATATTGAAA AAGGAAGAGT ATGAGTATTC AACATTTCCG 7320

TGTCGCCCTT ATTCCCTTTT TTGCGGCATT TTGCCTTCCT GTTTTTGCTC ACCCAGAAAC 7380

GCTGGTGAAA GTAAAAGATG CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAACT 7440

GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCCAATGAT 7500

GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC CGTATTGACG CCGGGCAAGA 7560

GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG GTTGAGTACT CACCAGTCAC 7620

AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA TGCAGTGCTG CCATAACCAT 7680

GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC 7740

CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT 7800

GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG CCTGTAGCAA TGGCAACAAC 7860

GTTGCGCAAA CTATTAACTG GCGAACTACT TACTCTAGCT TCCCGGCAAC AATTAATAGA 7920

CTGGATGGAG GCGGATAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG 7980

GTTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGTCT CGCGGTATCA TTGCAGCACT 8040

GGGGCCAGAT GGTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA GTCAGGCAAC 8100

TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA AGCATTGGTA 8160

ACTGTCAGAC CAAGTTTACT CATATATACT TTAGATTGAT TTAAAACTTC ATTTTTAATT 8220

TAAAAGGATC TAGGTGAAGA TCCTTTTTGA TAATCTCATG ACCAAAATCC CTTAACGTGA 8280

GTTTTCGTTC CACTGAGCGT CAGACCCCGT AGAAAAGATC AAAGGATCTT CTTGAGATCC 8340

TTTTTTTCTG CGCGTAATCT GCTGCTTGCA AACAAAAAAA CCACCGCTAC CAGCGGTGGT 8400

TTGTTTGCCG GATCAAGAGC TACCAACTCT TTTTCCGAAG GTAACTGGCT TCAGCAGAGC 8460

GCAGATACCA AATACTGTCC TTCTAGTGTA GCCGTAGTTA GGCCACCACT TCAAGAACTC 8520

TGTAGCACCG CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG CTGCCAGTGG 8580

CGATAAGTCG TGTCTTACCG GGTTGGACTC AAGACGATAG TTACCGGATA AGGCGCAGCG 8640

GTCGGGCTGA ACGGGGGGTT CGTGCACACA GCCCAGCTTG GAGCGAACGA CCTACACCGA 8700

ACTGAGATAC CTACAGCGTG AGCATTGAGA AAGCGCCACG CTTCCCGAAG GGAGAAAGGC 8760

GGACAGGTAT CCGGTAAGCG GCAGGGTCGG AACAGGAGAG CGCACGAGGG AGCTTCCAGG 8820

GGGAAACGCC TGGTATCTTT ATAGTCCTGT CGGGTTTCGC CACCTCTGAC TTGAGCGTCG 8880

ATTTTTGTGA TGCTCGTCAG GGGGGCGGAG CCTATGGAAA AACGCCAGCA ACGCGGCCTT 8940

TTTACGGTTC CTGGCCTTTT GCTGGCCTTT TGCTCACATG TTCTTTCCTG CGTTATCCCC 9000

TGATTCTGTG GATAACCGTA TTACCGCCTT TGAGTGAGCT GATACCGCTC GCCGCAGCCG 9060

AACGACCGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA GAGCGCCCAA TACGCAAACC 9120

GCCTCTCCCC GCGCGTTGGC CGATTCATTA ATGCAGCTGG CACGACAGGT TTCCCGACTG 9180

GAAAGCGGGC AGTGAGCGCA ACGCAATTAA TGTGAGTTAG CTCACTCATT AGGCACCCCA 9240

GGCTTTACAC TTTATGCTTC CGGCTCGTAT GTTGTGTGGA ATTGTGAGCG GATAACAATT 9300

TCACACAGGA AACAGCTATG ACCATGATTA CGCCA 9335

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGGGGATCCT CTAGACAATG ATATACATAG ATAAAAACC 39

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GGGAGATCTC CTTCGCTGTA CTATGTTATA AGAGAAGAG 39

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GGGGGATCCT GACTGCTTTG TCAAGGTTCA ATTCTGCTT 39

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GGGCCATGGA TCGCAGCCCT ACACATGTAA CAGTGTTGT 39

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

AAAGAGCTCT GAGGGCACTG AAGTCGCTTG ATGTGC 36

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GGGGAATTCT TGGATATATG CCGTGAACAA TTGTTATGTT AC 42

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5897 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AGCTTGCATG CCTGCAGATC TGCATGGGTG GAGACTTTTC AACAAAGGGT AATATCCGGA 60

AACCTCCTCG GATTCCATTG CCCAGCTATC TGTCACTTTA TTGTGAAGAT AGTGGAAAAG 120

GAAGGTGGCT CCTACAAATG CCATCATTGC GATAAAGGAA AGGCCATCGT TGAAGATGCC 180

TCTGCCGACA GTGGTCCCAA AGATGGACCC CCACCCACGA GGAGCATCGT GGAAAAAGAA 240

GACGTTCCAA CCACGTCTTC AAAGCAAGTG GATTGATGTG ATCATCGATG GAGACTTTTC 300

AACAAAGGGT AATATCCGGA AACCTCCTCG GATTCCATTG CCCAGCTATC TGTCACTTTA 360

TTGTGAAGAT AGTGGAAAAG GAAGGTGGCT CCTACAAATG CCATCATTGC GATAAAGGAA 420

AGGCCATCGT TGAAGATGCC TCTGCCGACA GTGGTCCCAA AGATGGACCC CCACCCACGA 480

GGAGCATCGT GGAAAAAGAA GACGTTCCAA CCACGTCTTC AAAGCAAGTG GATTGATGTG 540

ATATCTCCAC TGACGTAAGG GATGACGCAC AATCCCACTA TCCTTCGCAA GACCCTTCCT 600

CTATATAAGG AAGTTCATTT CATTTGGAGA GAACACGGGG GACTCTAGAG GATCCAGCTG 660

AAGGCTCGAC AAGGCAGTCC ACGGAGGAGC TGATATTTGG TGGACAAGCT GTGGATAGGA 720

GCAACCCTAT CCCTAATATA CCAGCACCAC CAAGTCAGGG CAATCCCCAG ATCAAGTGCA 780

AAGGTCCGCC TTGTTTCTCC TCTGTCTCTT GATCTGACTA ATCTTGGTTT ATGATTCGTT 840

GAGTAATTTT GGGGAAAGCT CCTTTGCTGC TCCACACATG TCCATTCGAA TTTTACCGTG 900

TTTAGCAAGG GCGAAAAGTT TGCATCTTGA TGATTTAGCT TGACTATGCG ATTGCTTTCC 960

TGGACCCGTG CAGCTGCGGA CGGATCTGGG GCCATTTGTT CCAGGCACGG GATAAGCATT 1020

CAGCCATGGT CCGTCCTGTA GAAACCCCAA CCCGTGAAAT CAAAAAACTC GACGGCCTGT 1080

GGGCATTCAG TCTGGATCGC GAAAACTGTG GAATTGATCA GCGTTGGTGG GAAAGCGCGT 1140

TACAAGAAAG CCGGGCAATT GCTGTGCCAG GCAGTTTTAA CGATCAGTTC GCCGATGCAG 1200

ATATTCGTAA TTATGCGGGC AACGTCTGGT ATCAGCGCGA AGTCTTTATA CCGAAAGGTT 1260

GGGCAGGCCA GCGTATCGTG CTGCGTTTCG ATGCGGTCAC TCATTACGGC AAAGTGTGGG 1320

TCAATAATCA GGAAGTGATG GAGCATCAGG GCGGCTATAC GCCATTTGAA GCCGATGTCA 1380

CGCCGTATGT TATTGCCGGG AAAAGTGTAC GTATCACCGT TTGTGTGAAC AACGAACTGA 1440

ACTGGCAGAC TATCCCGCCG GGAATGGTGA TTACCGACGA AAACGGCAAG AAAAAGCAGT 1500

CTTACTTCCA TGATTTCTTT AACTATGCCG GAATCCATCG CAGCGTAATG CTCTACACCA 1560

CGCCGAACAC CTGGGTGGAC GATATCACCG TGGTGACGCA TGTCGCGCAA GACTGTAACC 1620

ACGCGTCTGT TGACTGGCAG GTGGTGGCCA ATGGTGATGT CAGCGTTGAA CTGCGTGATG 1680

CGGATCAACA GGTGGTTGCA ACTGGACAAG GCACTAGCGG GACTTTGCAA GTGGTGAATC 1740

CGCACCTCTG GCAACCGGGT GAAGGTTATC TCTATGAACT GTGCGTCACA GCCAAAAGCC 1800

AGACAGAGTG TGATATCTAC CCGCTTCGCG TCGGCATCCG GTCAGTGGCA GTGAAGGGCG 1860

AACAGTTCCT GATTAACCAC AAACCGTTCT ACTTTACTGG CTTTGGTCGT CATGAAGATG 1920

CGGACTTACG TGGCAAAGGA TTCGATAACG TGCTGATGGT GCACGACCAC GCATTAATGG 1980

ACTGGATTGG GGCCAACTCC TACCGTACCT CGCATTACCC TTACGCTGAA GAGATGCTCG 2040

ACTGGGCAGA TGAACATGGC ATCGTGGTGA TTGATGAAAC TGCTGCTGTC GGCTTTAACC 2100

TCTCTTTAGG CATTGGTTTC GAAGCGGGCA ACAAGCCGAA AGAACTGTAC AGCGAAGAGG 2160

CAGTCAACGG GGAAACTCAG CAAGCGCACT TACAGGCGAT TAAAGAGCTG ATAGCGCGTG 2220

ACAAAAACCA CCCAAGCGTG GTGATGTGGA GTATTGCCAA CGAACCGGAT ACCCGTCCGC 2280

AAGTGCACGG GAATATTTCG CCACTGGCGG AAGCAACGCG TAAACTCGAC CCGACGCGTC 2340

CGATCACCTG CGTCAATGTA ATGTTCTGCG ACGCTCACAC CGATACCATC AGCGATCTCT 2400

TTGATGTGCT GTGCCTGAAC CGTTATTACG GATGGTATGT CCAAAGCGGC GATTTGGAAA 2460

CGGCAGAGAA GGTACTGGAA AAAGAACTTC TGGCCTGGCA GGAGAAACTG CATCAGCCGA 2520

TTATCATCAC CGAATACGGC GTGGATACGT TAGCCGGGCT GCACTCAATG TACACCGACA 2580

TGTGGAGTGA AGAGTATCAG TGTGCATGGC TGGATATGTA TCACCGCGTC TTTGATCGCG 2640

TCAGCGCCGT CGTCGGTGAA CAGGTATGGA ATTTCGCCGA TTTTGCGACC TCGCAAGGCA 2700

TATTGCGCGT TGGCGGTAAC AAGAAAGGGA TCTTCACTCG CGACCGCAAA CCGAAGTCGG 2760

CGGCTTTTCT GCTGCAAAAA CGCTGGACTG GCATGAACTT CGGTGAAAAA CCGCAGCAGG 2820

GAGGCAAACA ATGAATCAAC AACTCTCCTG GCGCACCATC GTCGGCTACA GCCTCGGTGG 2880

GGAATTGGAG AGCTCTGAGG GCACTGAAGT CGCTTGATGT GCTGAATTGT TTGTGATGTT 2940

GGTGGCGTAT TTTGTTTAAA TAAGTAAGCA TGGCTGTGAT TTTATCATAT GATCGATCTT 3000

TGGGGTTTTA TTTAACACAT TGTAAAATGT GTATCTATTA ATAACTCAAT GTATAAGATG 3060

TGTTCATTCT TCGGTTGCCA TAGATCTGCT TATTTGACCT GTGATGTTTT GACTCCAAAA 3120

ACCAAAATCA CAACTCAATA AACTCATGGA ATATGTCCAC CTGTTTCTTG AAGAGTTCAT 3180

CTACCATTCC AGTTGGCATT TATCAGTGTT GCAGCGGCGC TGTGCTTTGT AACATAACAA 3240

TTGTTCACGG CATATATCCA AGAATTCACT GGCCGTCGTT TTACAACGTC GTGACTGGGA 3300

AAACCCTGGC GTTACCCAAC TTAATCGCCT TGCAGCACAT CCCCCTTTCG CCAGCTGGCG 3360

TAATAGCGAA GAGGCCCGCA CCGATCGCCC TTCCCAACAG TTGCGCAGCC TGAATGGCGA 3420

ATGGCGCCTG ATGCGGTATT TTCTCCTTAC GCATCTGTGC GGTATTTCAC ACCGCATATG 3480

GTGCACTCTC AGTACAATCT GCTCTGATGC CGCATAGTTA AGCCAGCCCC GACACCCGCC 3540

AACACCCGCT GACGCGCCCT GACGGGCTTG TCTGCTCCCG GCATCCGCTT ACAGACAAGC 3600

TGTGACCGTC TCCGGGAGCT GCATGTGTCA GAGGTTTTCA CCGTCATCAC CGAAACGCGC 3660

GAGACGAAAG GGCCTCGTGA TACGCCTATT TTTATAGGTT AATGTCATGA TAATAATGGT 3720

TTCTTAGACG TCAGGTGGCA CTTTTCGGGG AAATGTGCGC GGAACCCCTA TTTGTTTATT 3780

TTTCTAAATA CATTCAAATA TGTATCCGCT CATGAGACAA TAACCCTGAT AAATGCTTCA 3840

ATAATATTGA AAAAGGAAGA GTATGAGTAT TCAACATTTC CGTGTCGCCC TTATTCCCTT 3900

TTTTGCGGCA TTTTGCCTTC CTGTTTTTGC TCACCCAGAA ACGCTGGTGA AAGTAAAAGA 3960

TGCTGAAGAT CAGTTGGGTG CACGAGTGGG TTACATCGAA CTGGATCTCA ACAGCGGTAA 4020

GATCCTTGAG AGTTTTCGCC CCGAAGAACG TTTTCCAATG ATGAGCACTT TTAAAGTTCT 4080

GCTATGTGGC GCGGTATTAT CCCGTATTGA CGCCGGGCAA GAGCAACTCG GTCGCCGCAT 4140

ACACTATTCT CAGAATGACT TGGTTGAGTA CTCACCAGTC ACAGAAAAGC ATCTTACGGA 4200

TGGCATGACA GTAAGAGAAT TATGCAGTGC TGCCATAACC ATGAGTGATA ACACTGCGGC 4260

CAACTTACTT CTGACAACGA TCGGAGGACC GAAGGAGCTA ACCGCTTTTT TGCACAACAT 4320

GGGGGATCAT GTAACTCGCC TTGATCGTTG GGAACCGGAG CTGAATGAAG CCATACCAAA 4380

CGACGAGCGT GACACCACGA TGCCTGTAGC AATGGCAACA ACGTTGCGCA AACTATTAAC 4440

TGGCGAACTA CTTACTCTAG CTTCCCGGCA ACAATTAATA GACTGGATGG AGGCGGATAA 4500

AGTTGCAGGA CCACTTCTGC GCTCGGCCCT TCCGGCTGGC TGGTTTATTG CTGATAAATC 4560

TGGAGCCGGT GAGCGTGGGT CTCGCGGTAT CATTGCAGCA CTGGGGCCAG ATGGTAAGCC 4620

CTCCCGTATC GTAGTTATCT ACACGACGGG GAGTCAGGCA ACTATGGATG AACGAAATAG 4680

ACAGATCGCT GAGATAGGTG CCTCACTGAT TAAGCATTGG TAACTGTCAG ACCAAGTTTA 4740

CTCATATATA CTTTAGATTG ATTTAAAACT TCATTTTTAA TTTAAAAGGA TCTAGGTGAA 4800

GATCCTTTTT GATAATCTCA TGACCAAAAT CCCTTAACGT GAGTTTTCGT TCCACTGAGC 4860

GTCAGACCCC GTAGAAAAGA TCAAAGGATC TTCTTGAGAT CCTTTTTTTC TGCGCGTAAT 4920

CTGCTGCTTG CAAACAAAAA AACCACCGCT ACCAGCGGTG GTTTGTTTGC CGGATCAAGA 4980

GCTACCAACT CTTTTTCCGA AGGTAACTGG CTTCAGCAGA GCGCAGATAC CAAATACTGT 5040

CCTTCTAGTG TAGCCGTAGT TAGGCCACCA CTTCAAGAAC TCTGTAGCAC CGCCTACATA 5100

CCTCGCTCTG CTAATCCTGT TACCAGTGGC TGCTGCCAGT GGCGATAAGT CGTGTCTTAC 5160

CGGGTTGGAC TCAAGACGAT AGTTACCGGA TAAGGCGCAG CGGTCGGGCT GAACGGGGGG 5220

TTCGTGCACA CAGCCCAGCT TGGAGCGAAC GACCTACACC GAACTGAGAT ACCTACAGCG 5280

TGAGCATTGA GAAAGCGCCA CGCTTCCCGA AGGGAGAAAG GCGGACAGGT ATCCGGTAAG 5340

CGGCAGGGTC GGAACAGGAG AGCGCACGAG GGAGCTTCCA GGGGGAAACG CCTGGTATCT 5400

TTATAGTCCT GTCGGGTTTC GCCACCTCTG ACTTGAGCGT CGATTTTTGT GATGCTCGTC 5460

AGGGGGGCGG AGCCTATGGA AAAACGCCAG CAACGCGGCC TTTTTACGGT TCCTGGCCTT 5520

TTGCTGGCCT TTTGCTCACA TGTTCTTTCC TGCGTTATCC CCTGATTCTG TGGATAACCG 5580

TATTACCGCC TTTGAGTGAG CTGATACCGC TCGCCGCAGC CGAACGACCG AGCGCAGCGA 5640

GTCAGTGAGC GAGGAAGCGG AAGAGCGCCC AATACGCAAA CCGCCTCTCC CCGCGCGTTG 5700

GCCGATTCAT TAATGCAGCT GGCACGACAG GTTTCCCGAC TGGAAAGCGG GCAGTGAGCG 5760

CAACGCAATT AATGTGAGTT AGCTCACTCA TTAGGCACCC CAGGCTTTAC ACTTTATGCT 5820

TCCGGCTCGT ATGTTGTGTG GAATTGTGAG CGGATAACAA TTTCACACAG GAAACAGCTA 5880

TGACCATGAT TACGCCA 5897

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6898 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

AGCTTGCATG CCTGCAGTGC AGCGTGACCC GGTCGTGCCC CTCTCTAGAG ATAATGAGCA 60

TTGCATGTCT AAGTTATAAA AAATTACCAC ATATTTTTTT TGTCACACTT GTTTGAAGTG 120

CAGTTTATCT ATCTTTATAC ATATATTTAA ACTTTAATCT ACGAATAATA TAATCTATAG 180

TACTACAATA ATATCAGTGT TTTAGAGAAT CATATAAATG AACAGTTAGA CATGGTCTAA 240

AGGACAATTG AGTATTTTGA CAACAGGACT CTACAGTTTT ATCTTTTTAG TGTGCATGTG 300

TTCTCCTTTT TTTTTGCAAA TAGCTTCACC TATATAATAC TTCATCCATT TTATTAGTAC 360

ATCCATTTAG GGTTTAGGGT TAATGGTTTT TATAGACTAA TTTTTTTAGT ACATCTATTT 420

TATTCTATTT TAGCCTCTAA ATTAAGAAAA CTAAAACTCT ATTTTAGTTT TTTTATTTAA 480

TAATTTAGAT ATAAAATAGA ATAAAATAAA GTGACTAAAA ATTAAACAAA TACCCTTTAA 540

GAAATTAAAA AAACTAAGGA AACATTTTTC TTGTTTCGAG TAGATAATGC CAGCCTGTTA 600

AACGCCGTCG ACGAGTCTAA CGGACACCAA CCAGCGAACC AGCAGCGTCG CGTCGGGCCA 660

AGCGAAGCAG ACGGCACGGC ATCTCTGTCG CTGCCTCTGG ACCCCTCTCG AGAGTTCCGC 720

TCCACCGTTG GACTTGCTCC GCTGTCGGCA TCCAGAAATT GCGTGGCGGA GCGGCAGACG 780

TGAGCCGGCA CGGCAGGCGG CCTCCTCCTC CTCTCACGGC ACGGCAGCTA CGGGGGATTC 840

CTTTCCCACC GCTCCTTCGC TTTCCCTTCC TCGCCCGCCG TAATAAATAG ACACCCCCTC 900

CACACCCTCT TTCCCCAACC TCGTGTTGTT CGGAGCGCAC ACACACACAA CCAGATCTCC 960

CCCAAATCCA CCCGTCGGCA CCTCCGCTTC AAGGTACGCC GCTCGTCCTC CCCCCCCCCC 1020

CCTCTCTACC TTCTCTAGAT CGGCGTTCCG GTCCATGCAT GGTTAGGGCC CGGTAGTTCT 1080

ACTTCTGTTC ATGTTTGTGT TAGATCCGTG TTTGTGTTAG ATCCGTGCTG CTAGCGTTCG 1140

TACACGGATG CGACCTGTAC GTCAGACACG TTCTGATTGC TAACTTGCCA GTGTTTCTCT 1200

TTGGGGAATC CTGGGATGGC TCTAGCCGTT CCGCAGACGG GATCGATTTC ATGATTTTTT 1260

TTGTTTCGTT GCATAGGGTT TGGTTTGCCC TTTTCCTTTA TTTCAATATA TGCCGTGCAC 1320

TTGTTTGTCG GGTCATCTTT TCATGCTTTT TTTTGTCTTG GTTGTGATGA TGTGGTCTGG 1380

TTGGGCGGTC GTTCTAGATC GGAGTAGAAT TCTGTTTCAA ACTACCTGGT GGATTTATTA 1440

ATTTTGGATC TGTATGTGTG TGCCATACAT ATTCATAGTT ACGAATTGAA GATGATGGAT 1500

GGAAATATCG ATCTAGGATA GGTATACATG TTGATGCGGG TTTTACTGAT GCATATACAG 1560

AGATGCTTTT TGTTCGCTTG GTTGTGATGA TGTGGTGTGG TTGGGCGGTC GTTCATTCGT 1620

TCTAGATCGG AGTAGAATAC TGTTTCAAAC TACCTGGTGT ATTTATTAAT TTTGGAACTG 1680

TATGTGTGTG TCATACATCT TCATAGTTAC GAGTTTAAGA TGGATGGAAA TATCGATCTA 1740

GGATAGGTAT ACATGTTGAT GTGGGTTTTA CTGATGCATA TACATGATGG CATATGCAGC 1800

ATCTATTCAT ATGCTCTAAC CTTGAGTACC TATCTATTAT AATAAACAAG TATGTTTTAT 1860

AATTATTTTG ATCTTGATAT ACTTGGATGA TGGCATATGC AGCAGCTATA TGTGGATTTT 1920

TTTAGCCCTG CCTTCATACG CTATTTATTT GCTTGGTACT GTTTCTTTTG TCGATGCTCA 1980

CCCTGTTGTT TGGTGTTACT TCTGCAGGGT ACCCCCGGGG TCGACCATGG TCCGTCCTGT 2040

AGAAACCCCA ACCCGTGAAA TCAAAAAACT CGACGGCCTG TGGGCATTCA GTCTGGATCG 2100

CGAAAACTGT GGAATTGATC AGCGTTGGTG GGAAAGCGCG TTACAAGAAA GCCGGGCAAT 2160

TGCTGTGCCA GGCAGTTTTA ACGATCAGTT CGCCGATGCA GATATTCGTA ATTATGCGGG 2220

CAACGTCTGG TATCAGCGCG AAGTCTTTAT ACCGAAAGGT TGGGCAGGCC AGCGTATCGT 2280

GCTGCGTTTC GATGCGGTCA CTCATTACGG CAAAGTGTGG GTCAATAATC AGGAAGTGAT 2340

GGAGCATCAG GGCGGCTATA CGCCATTTGA AGCCGATGTC ACGCCGTATG TTATTGCCGG 2400

GAAAAGTGTA CGTATCACCG TTTGTGTGAA CAACGAACTG AACTGGCAGA CTATCCCGCC 2460

GGGAATGGTG ATTACCGACG AAAACGGCAA GAAAAAGCAG TCTTACTTCC ATGATTTCTT 2520

TAACTATGCC GGAATCCATC GCAGCGTAAT GCTCTACACC ACGCCGAACA CCTGGGTGGA 2580

CGATATCACC GTGGTGACGC ATGTCGCGCA AGACTGTAAC CACGCGTCTG TTGACTGGCA 2640

GGTGGTGGCC AATGGTGATG TCAGCGTTGA ACTGCGTGAT GCGGATCAAC AGGTGGTTGC 2700

AACTGGACAA GGCACTAGCG GGACTTTGCA AGTGGTGAAT CCGCACCTCT GGCAACCGGG 2760

TGAAGGTTAT CTCTATGAAC TGTGCGTCAC AGCCAAAAGC CAGACAGAGT GTGATATCTA 2820

CCCGCTTCGC GTCGGCATCC GGTCAGTGGC AGTGAAGGGC GAACAGTTCC TGATTAACCA 2880

CAAACCGTTC TACTTTACTG GCTTTGGTCG TCATGAAGAT GCGGACTTAC GTGGCAAAGG 2940

ATTCGATAAC GTGCTGATGG TGCACGACCA CGCATTAATG GACTGGATTG GGGCCAACTC 3000

CTACCGTACC TCGCATTACC CTTACGCTGA AGAGATGCTC GACTGGGCAG ATGAACATGG 3060

CATCGTGGTG ATTGATGAAA CTGCTGCTGT CGGCTTTAAC CTCTCTTTAG GCATTGGTTT 3120

CGAAGCGGGC AACAAGCCGA AAGAACTGTA CAGCGAAGAG GCAGTCAACG GGGAAACTCA 3180

GCAAGCGCAC TTACAGGCGA TTAAAGAGCT GATAGCGCGT GACAAAAACC ACCCAAGCGT 3240

GGTGATGTGG AGTATTGCCA ACGAACCGGA TACCCGTCCG CAAGTGCACG GGAATATTTC 3300

GCCACTGGCG GAAGCAACGC GTAAACTCGA CCCGACGCGT CCGATCACCT GCGTCAATGT 3360

AATGTTCTGC GACGCTCACA CCGATACCAT CAGCGATCTC TTTGATGTGC TGTGCCTGAA 3420

CCGTTATTAC GGATGGTATG TCCAAAGCGG CGATTTGGAA ACGGCAGAGA AGGTACTGGA 3480

AAAAGAACTT CTGGCCTGGC AGGAGAAACT GCATCAGCCG ATTATCATCA CCGAATACGG 3540

CGTGGATACG TTAGCCGGGC TGCACTCAAT GTACACCGAC ATGTGGAGTG AAGAGTATCA 3600

GTGTGCATGG CTGGATATGT ATCACCGCGT CTTTGATCGC GTCAGCGCCG TCGTCGGTGA 3660

ACAGGTATGG AATTTCGCCG ATTTTGCGAC CTCGCAAGGC ATATTGCGCG TTGGCGGTAA 3720

CAAGAAAGGG ATCTTCACTC GCGACCGCAA ACCGAAGTCG GCGGCTTTTC TGCTGCAAAA 3780

ACGCTGGACT GGCATGAACT TCGGTGAAAA ACCGCAGCAG GGAGGCAAAC AATGAATCAA 3840

CAACTCTCCT GGCGCACCAT CGTCGGCTAC AGCCTCGGTG GGGAATTGGA GAGCTCTGAG 3900

GGCACTGAAG TCGCTTGATG TGCTGAATTG TTTGTGATGT TGGTGGCGTA TTTTGTTTAA 3960

ATAAGTAAGC ATGGCTGTGA TTTTATCATA TGATCGATCT TTGGGGTTTT ATTTAACACA 4020

TTGTAAAATG TGTATCTATT AATAACTCAA TGTATAAGAT GTGTTCATTC TTCGGTTGCC 4080

ATAGATCTGC TTATTTGACC TGTGATGTTT TGACTCCAAA AACCAAAATC ACAACTCAAT 4140

AAACTCATGG AATATGTCCA CCTGTTTCTT GAAGAGTTCA TCTACCATTC CAGTTGGCAT 4200

TTATCAGTGT TGCAGCGGCG CTGTGCTTTG TAACATAACA ATTGTTCACG GCATATATCC 4260

AAGAATTCAC TGGCCGTCGT TTTACAACGT CGTGACTGGG AAAACCCTGG CGTTACCCAA 4320

CTTAATCGCC TTGCAGCACA TCCCCCTTTC GCCAGCTGGC GTAATAGCGA AGAGGCCCGC 4380

ACCGATCGCC CTTCCCAACA GTTGCGCAGC CTGAATGGCG AATGGCGCCT GATGCGGTAT 4440

TTTCTCCTTA CGCATCTGTG CGGTATTTCA CACCGCATAT GGTGCACTCT CAGTACAATC 4500

TGCTCTGATG CCGCATAGTT AAGCCAGCCC CGACACCCGC CAACACCCGC TGACGCGCCC 4560

TGACGGGCTT GTCTGCTCCC GGCATCCGCT TACAGACAAG CTGTGACCGT CTCCGGGAGC 4620

TGCATGTGTC AGAGGTTTTC ACCGTCATCA CCGAAACGCG CGAGACGAAA GGGCCTCGTG 4680

ATACGCCTAT TTTTATAGGT TAATGTCATG ATAATAATGG TTTCTTAGAC GTCAGGTGGC 4740

ACTTTTCGGG GAAATGTGCG CGGAACCCCT ATTTGTTTAT TTTTCTAAAT ACATTCAAAT 4800

ATGTATCCGC TCATGAGACA ATAACCCTGA TAAATGCTTC AATAATATTG AAAAAGGAAG 4860

AGTATGAGTA TTCAACATTT CCGTGTCGCC CTTATTCCCT TTTTTGCGGC ATTTTGCCTT 4920

CCTGTTTTTG CTCACCCAGA AACGCTGGTG AAAGTAAAAG ATGCTGAAGA TCAGTTGGGT 4980

GCACGAGTGG GTTACATCGA ACTGGATCTC AACAGCGGTA AGATCCTTGA GAGTTTTCGC 5040

CCCGAAGAAC GTTTTCCAAT GATGAGCACT TTTAAAGTTC TGCTATGTGG CGCGGTATTA 5100

TCCCGTATTG ACGCCGGGCA AGAGCAACTC GGTCGCCGCA TACACTATTC TCAGAATGAC 5160

TTGGTTGAGT ACTCACCAGT CACAGAAAAG CATCTTACGG ATGGCATGAC AGTAAGAGAA 5220

TTATGCAGTG CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT TCTGACAACG 5280

ATCGGAGGAC CGAAGGAGCT AACCGCTTTT TTGCACAACA TGGGGGATCA TGTAACTCGC 5340

CTTGATCGTT GGGAACCGGA GCTGAATGAA GCCATACCAA ACGACGAGCG TGACACCACG 5400

ATGCCTGTAG CAATGGCAAC AACGTTGCGC AAACTATTAA CTGGCGAACT ACTTACTCTA 5460

GCTTCCCGGC AACAATTAAT AGACTGGATG GAGGCGGATA AAGTTGCAGG ACCACTTCTG 5520

CGCTCGGCCC TTCCGGCTGG CTGGTTTATT GCTGATAAAT CTGGAGCCGG TGAGCGTGGG 5580

TCTCGCGGTA TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT CGTAGTTATC 5640

TACACGACGG GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC TGAGATAGGT 5700

GCCTCACTGA TTAAGCATTG GTAACTGTCA GACCAAGTTT ACTCATATAT ACTTTAGATT 5760

GATTTAAAAC TTCATTTTTA ATTTAAAAGG ATCTAGGTGA AGATCCTTTT TGATAATCTC 5820

ATGACCAAAA TCCCTTAACG TGAGTTTTCG TTCCACTGAG CGTCAGACCC CGTAGAAAAG 5880

ATCAAAGGAT CTTCTTGAGA TCCTTTTTTT CTGCGCGTAA TCTGCTGCTT GCAAACAAAA 5940

AAACCACCGC TACCAGCGGT GGTTTGTTTG CCGGATCAAG AGCTACCAAC TCTTTTTCCG 6000

AAGGTAACTG GCTTCAGCAG AGCGCAGATA CCAAATACTG TCCTTCTAGT GTAGCCGTAG 6060

TTAGGCCACC ACTTCAAGAA CTCTGTAGCA CCGCCTACAT ACCTCGCTCT GCTAATCCTG 6120

TTACCAGTGG CTGCTGCCAG TGGCGATAAG TCGTGTCTTA CCGGGTTGGA CTCAAGACGA 6180

TAGTTACCGG ATAAGGCGCA GCGGTCGGGC TGAACGGGGG GTTCGTGCAC ACAGCCCAGC 6240

TTGGAGCGAA CGACCTACAC CGAACTGAGA TACCTACAGC GTGAGCATTG AGAAAGCGCC 6300

ACGCTTCCCG AAGGGAGAAA GGCGGACAGG TATCCGGTAA GCGGCAGGGT CGGAACAGGA 6360

GAGCGCACGA GGGAGCTTCC AGGGGGAAAC GCCTGGTATC TTTATAGTCC TGTCGGGTTT 6420

CGCCACCTCT GACTTGAGCG TCGATTTTTG TGATGCTCGT CAGGGGGGCG GAGCCTATGG 6480

AAAAACGCCA GCAACGCGGC CTTTTTACGG TTCCTGGCCT TTTGCTGGCC TTTTGCTCAC 6540

ATGTTCTTTC CTGCGTTATC CCCTGATTCT GTGGATAACC GTATTACCGC CTTTGAGTGA 6600

GCTGATACCG CTCGCCGCAG CCGAACGACC GAGCGCAGCG AGTCAGTGAG CGAGGAAGCG 6660

GAAGAGCGCC CAATACGCAA ACCGCCTCTC CCCGCGCGTT GGCCGATTCA TTAATGCAGC 6720

TGGCACGACA GGTTTCCCGA CTGGAAAGCG GGCAGTGAGC GCAACGCAAT TAATGTGAGT 6780

TAGCTCACTC ATTAGGCACC CCAGGCTTTA CACTTTATGC TTCCGGCTCG TATGTTGTGT 6840

GGAATTGTGA GCGGATAACA ATTTCACACA GGAAACAGCT ATGACCATGA TTACGCCA 6898

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CAGATCTGCA GATCTGCATG GGCGATG 27

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GGGGACTCTA GAGGATCCCC GGGTGGTCAG TCCCTT 36

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GAATTTCCCC 10

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GATCCGGATC CG 12

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

TCGACGGATC CG 12

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GGGGACTCTA GAGGATCCCG AATTTCCCC 29

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 57 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GATCCAGCTG AAGGCTCGAC AAGGCAGATC CACGGAGGAG CTGATATTTG GTGGACA 57

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 57 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

AGCTTGTCCA CCAAATATCA GCTCCTCCGT GGATCTGCCT TGTCCAGCCT TCAGCTG 57

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 64 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

AGCTGTGGAT AGGAGCAACC CTATCCCTAA TATACCAGCA CCACCAAGTC AGGGCAATCC 60

CGGG 64

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 64 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

TCGACCCGGG ATTGCCCTGA CTTGGTGGTG CTGGTATATT AGGGATAGGG TTGCTCCTAT 60

CCAC 64

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 62 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

CCGGGCCATT TGTTCCAGGC ACGGGATAAG CATTCAGCCA TGGGATATCA AGCTTGGATC 60

CC 62

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 62 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

TCGAGGGATC CAAGCTTGAT ATCCCATGGC TGAATGCTTA TCCCGTGCCT GGAACAAATG 60

GC 62

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GATATCAAGC TTGGATCCC 19

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CGGTACCTCG AGTTAAC 17

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

CATGGTTAAC TCGAGGTACC GAGCT 25

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

ATCTGCATGG GTG 13

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GGGGACTCTA GAGGATCCAG 20

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

GTTAACTCGA GGTACCGAGC TCGAATTTCC CC 32

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

GAGTTCAGGC TTTTTCATAG CT 22

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

AGATCTCGTG AGATAATGAA AAAG 24

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

ACTCGCCGAT AGTGGAAACC GACGCCCCAG CACTCGTCCG AGGGCAAAGG AATAGTAAGA 60

GCTCGG 66

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 70 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

GATCCCGAGC TCTTACTATT CCTTTGCCCT CGGACGAGTG CTGGGGCGTC GGTTTCCACT 60

ATCGGCGAGT 70

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 88 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

CTGCAGGCCG GCCTTAATTA AGCGGCCGCG TTTAAACGCC CGGGCATTTA AATGGCGCGC 60

CGCGATCGCT TGCAGATCTG CATGGGTG 88

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

GACGGATCTG 10

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

TGAGATCTGA GCTCGAATTT CCCC 24

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

GGTACCCCCG GGGTCGACCA TGG 23

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GGGAATTGGA GCTCGAATTT CCCC 24

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

GGGAAATTAA GCTT 14

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 69 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

AGCGGCCGCA TTCCCGGGAA GCTTGCATGC CTGCAGAGAT CCGGTACCCG GGGATCCTCT 60
AGAGTCGAC 69

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GGTACCCCCG GGGTCGACCA TGGTTAACTC GAGGTACCGA GCTCGAATTT CCCC 54

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

GGGAATTGGT TTAAACGCGG CCGCTT 26

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

CCATGCATGG 10

50476

148